REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
February 15, 1995

3. REPORT TYPE AND DATES COVERED
Final - July 1, 1992 to Dec. 31, 1994

4. TITLE AND SUBTITLE
New Adaptive Methods for Reconfigurable Flight Control Systems -- Appendix II (of II)

5. FUNDING NUMBERS
F49 620-92-J-0386
61102F
2304/AS

6. AUTHOR(S)
Marc Bodson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Carnegie Mellon University
500 Forbes Avenue
Pittsburgh, PA 15213-3890

8. PERFORMING ORGANIZATION REPORT NUMBER
AEOSR-TR- 95 - 0231

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
Bolling, AFB DC 20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-92-J-0386

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The report discusses methods for the design of reconfigurable flight control systems. Model reference adaptive control algorithms were developed which allow for the rapid identification of aircraft parameters after failures or damages and for automatic control redesign. These algorithms are simple enough that they can be implemented in real-time with existing computers. Simulation results were obtained using a detailed nonlinear model of a fighter aircraft. The results demonstrate the ability of the adaptive algorithms to automatically readjust trim values after failures, to restore tracking of the pilot commands despite the loss of actuator effectiveness, and to coordinate the use of the remaining surfaces in order to maintain the decoupling between the rotational axes. The report also discusses analytical results that were obtained during the course of the research regarding the stability and convergence properties of multivariable adaptive control algorithms, the use of averaging methods for the analysis of transient response and robustness, and the estimation of uncertainty in dynamic models. A modified recursive least-squares algorithm with forgetting factor is also described which gives fast adaptation under normal conditions while keeping the sensitivity to noise limited when signal to noise ratios are poor.

14. SUBJECT TERMS
Flight Control, reconfigurable control, multivariable systems, robust control, adaptive control

15. NUMBER OF PAGES
179

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)

Appendix II

Multivariable Adaptive Control
Reduced Prior Information, Convergence, and Stability

Michel de Mathelin

Electrical & Computer Engineering Department
Carnegie Mellon University
Pittsburgh PA 15213

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

January 15, 1993

Advisor: Professor Marc Bodson

Acknowledgements

My first thanks go to Professor Marc Bodson, my advisor and also a friend. He had a crucial role in my decision to undertake a Ph.D. at Carnegie Mellon. To accomplish my Ph.D. studies under his guidance was an extremely formative and enjoyable experience. His enthusiasm and constant support made this thesis possible. His dedication to academic excellence will have a long lasting effect on me.

I also thank Professors Bruce Krogh, José Moura, and Pradeep Khosla at Carnegie Mellon and Professor John Chiasson at the University of Pittsburgh for numerous and fruitful interactions during these past five years. I am particularly indebted to Professors Bruce Krogh, José Moura, and John Chiasson who had the strenuous task of reviewing this thesis.

I am particularly thankful to friends at Carnegie Mellon with whom I interacted thorough these years. Especially, Nikos Papanikolopoulos, Chris Paredis, Peter Feldman, and Larry Holloway. I would also like to mention fellow Ph.D. students Jacques Benkoski, Wayne Carriker, Radu Jasinschi, Brad Nelson, Sandy Ramos, Jim Regh, Alexei Sacks, R.S. Sreenivas, Jennifer Stephan, Matt Vea, and Reg Willson. Some of them have graduated now and I wish them good luck in their new environment. Special thanks are due to Debbie Scappatura who was always ready to help when her extensive secretarial skills were needed. I would like to express my gratitude to Judy Bandola, Pat Guy, Elaine Lawrence, Carol Patterson, and Lynn Philibin in the Electrical and Computer Engineering Department at Carnegie Mellon for their administrative support.

Part of my work was done at Louvain University thanks to the hospitality of Professors Michel Gevers, George Bastin, and Guy Campion. I am extremely thankful to them for their advice and their support. They also were instrumental in my decision to go to graduate school. I wish to express my sincere appreciation to Vincent Blondel, Libei Chen, Marc De Neyer, Daniele Galardini, Marc Haest, and Philippe Lefevre for numerous and interesting interactions during my stay at Louvain University.

I am also grateful for the logistical support received from Christiaan Perneel and Professor Marc Acheroy at the Royal Military Academy in Brussels. I am also very appreciative of their efforts to make my stay at the Royal Military Academy as pleasant and as fruitful as possible.

I acknowledge the support of several organizations, including the Belgian American Educational Foundation, the National Science Foundation (Grant ECS 88-10145 and CISE 88-20575),

the Air Force Office of Scientific Research (Grant F49620-92-J-0386), and the Electrical and Computer Engineering Department at Carnegie Mellon University.

Finally, I dedicate this thesis to my wife Emmanuelle whose constant love and support made my Ph.D. studies possible.

Contents

List of Tables

List of Figures

Symbols and acronyms

block diag$\{X_i\}$	Block diagonal matrix with i-th matrix diagonal element X_i.
deg	Degree of a polynomial.
det	Determinant of a matrix.
diag$\{x_i\}$	Diagonal matrix with i-th scalar diagonal elements x_i.
$H(s)$	Hermite normal form.
$h^{(i)}$	i-th Markov parameter.
I	Identity matrix of unspecified dimension.
I_p	Identity matrix of dimension p.
K_p	High-frequency gain matrix.
LQ	Linear Quadratic.
LQG	Linear Quadratic Gaussian.
MFD	Matrix fraction description.
MIMO	Multi-input, multi-output.
MRAC	Model reference adaptive control.
n	Order of system.
PE	Persistently exciting.
PPAC	Adaptive pole placement.
$P_{\mathcal{X}}$	Orthogonal projection onto subspace $\mathcal{X}$.
$P_{\mathcal{X}}^{\perp}$	Orthogonal projection onto subspace $\mathcal{X}^{\perp}$, the orthogonal complement of $\mathcal{X}$.
$\Re^{p\times m}$	The set of real matrices with p rows and m columns.
$\Re^{p\times m}(s)$	The set of $(p \times m)$ matrices whose elements are rational functions of s.
$\Re^{p\times m}[s]$	The set of $(p \times m)$ matrices whose elements are polynomials in s.
SE	Sufficiently exciting.
SISO	Single-input, single-output.
SPR	Strictly positive real.

STR	Self-tuning regulator.
∂X	Degree of polynomial X or largest degree of polynomial elements of X.
$\partial c_i X$	Largest degree of polynomial elements of i-th column of X.
$\partial r_i X$	Largest degree of polynomial elements of i-th row of X.
$\Gamma_c[X]$	Matrix of the leading coefficients of the polynomial elements of X whose degree is the largest in their column.
$\Gamma_r[X]$	Matrix of the leading coefficients of the polynomial elements of X whose degree is the largest in their row.
κ_i	i-th pseudo-controllability index.
$\kappa_{\max}$	Maximum of the pseudo-controllability indices.
$\lambda_i(X)$	i-th eigenvalue of X.
$\lambda_{\max}(X), \lambda_{\min}(X)$	Largest and smallest eigenvalues of X.
μ_i	i-th controllability index.
$\mu_{\max}$	Maximum of the controllability indices.
ν_i	i-th observability index.
$\nu_{\max}$	Maximum of the observability indices.
ρ_i	i-th pseudo-observability index.
$\rho_{\max}$	Maximum of the pseudo-observability indices.
$\sigma_i(X)$	i-th singular value of X.
$\sigma_{\max}(X), \sigma_{\min}(X)$	Largest and smallest singular values of X.

Introduction

Standard control design techniques are generally based on the knowledge of a good model of the system. When such knowledge is missing, for example when the system is unknown or time-varying, it seems intuitive to combine system identification techniques with control methods. Adaptive control combines both approaches and adapts the control law to the unknown and possibly time-varying environment. In this work, we are mostly concerned with *parametric* estimation and *parametric* adaptive control of linear systems, where the system is parameterized and only parameters need to be determined. Parametric adaptive control techniques for single-input, single-output (SISO) systems assume some kind of *a priori* knowledge about the system. Usually, this *a priori* knowledge is about the structure of the system like, *e.g.*, the order of the system. The design of the adaptive controller is then based on this *a priori* structural knowledge. A typical adaptive control system will consist of a control law, a parameter identifier using measurements of the input and the output of the system, and an adjustment mechanism to update the parameters of the controller based on the identification. A general representation of an adaptive control system is shown in Fig. 0.1.

Adaptive control techniques for SISO systems have been extended to multivariable (multi-input, multi-output (MIMO)) systems. Similarly to the SISO case, MIMO parametric adaptive control schemes require *a priori* information about the structure of the system. However, in the MIMO case, this required *a priori* information can be so elaborate that it greatly reduces the overall applicability of current MIMO adaptive control techniques. Furthermore, current MIMO adaptive control algorithms often do guarantee neither stability nor parameter convergence. The main objective of this thesis is to address these three important issues: *a priori* information, stability, and parameter convergence in MIMO adaptive control.

0.1 Overview of adaptive control

In the past years, several methods have been developed for designing adaptive controllers. These techniques can be separated in two different classes: indirect and direct adaptive control. Indirect adaptive control is the most straightforward approach. It consists in estimating the parameters of the transfer function of the process or the parameters of a state-space representation of the process (and parameters of the disturbances when stochastic models are used).

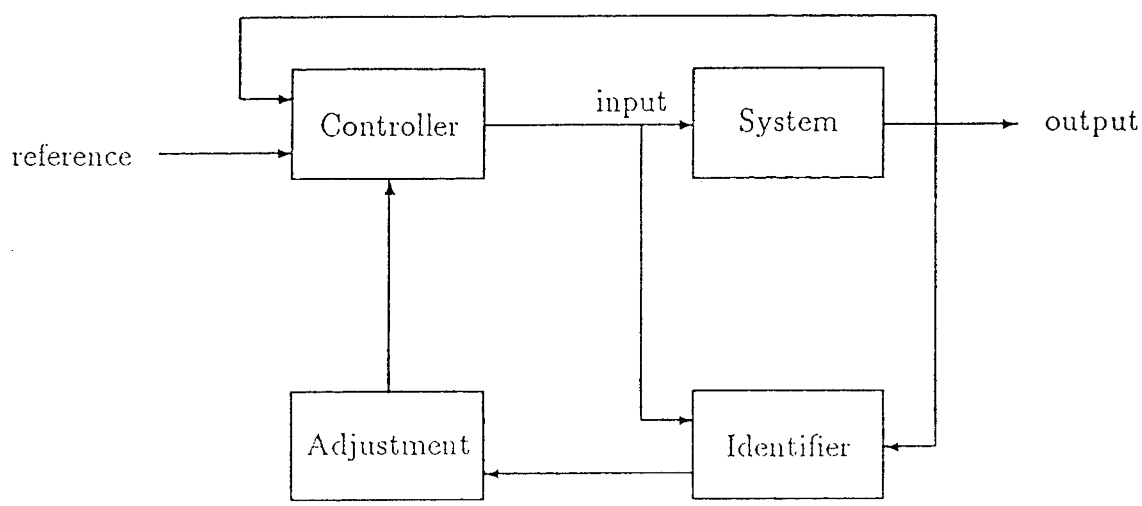

Figure 0.1: Adaptive control system

Then, the controller is updated based on the estimates of the parameters using an adjustment mechanism corresponding to some control design method (*cf.* Fig 0.1). The design of the controller is independent of the estimation. Any standard design technique could potentially be considered. The adjustment mechanism is often a complex nonlinear transformation between the estimated parameters and the controller parameters. In contrast, in direct adaptive control, the parameters of the controller are estimated directly. In that case, the estimator updates directly the regulator parameters and the adjustment mechanism in the adaptive control system is simply the identity transformation.

It is possible to classify further the different adaptive control techniques by the type of control law used. A possible classification of the main SISO adaptive controller design methods is presented in Fig. 0.2. A well-known technique is the direct or indirect *model reference adaptive control* (MRAC). In this scheme, the desired output of the closed-loop system is specified through a reference model, and the adaptive controller tries to make the plant output match

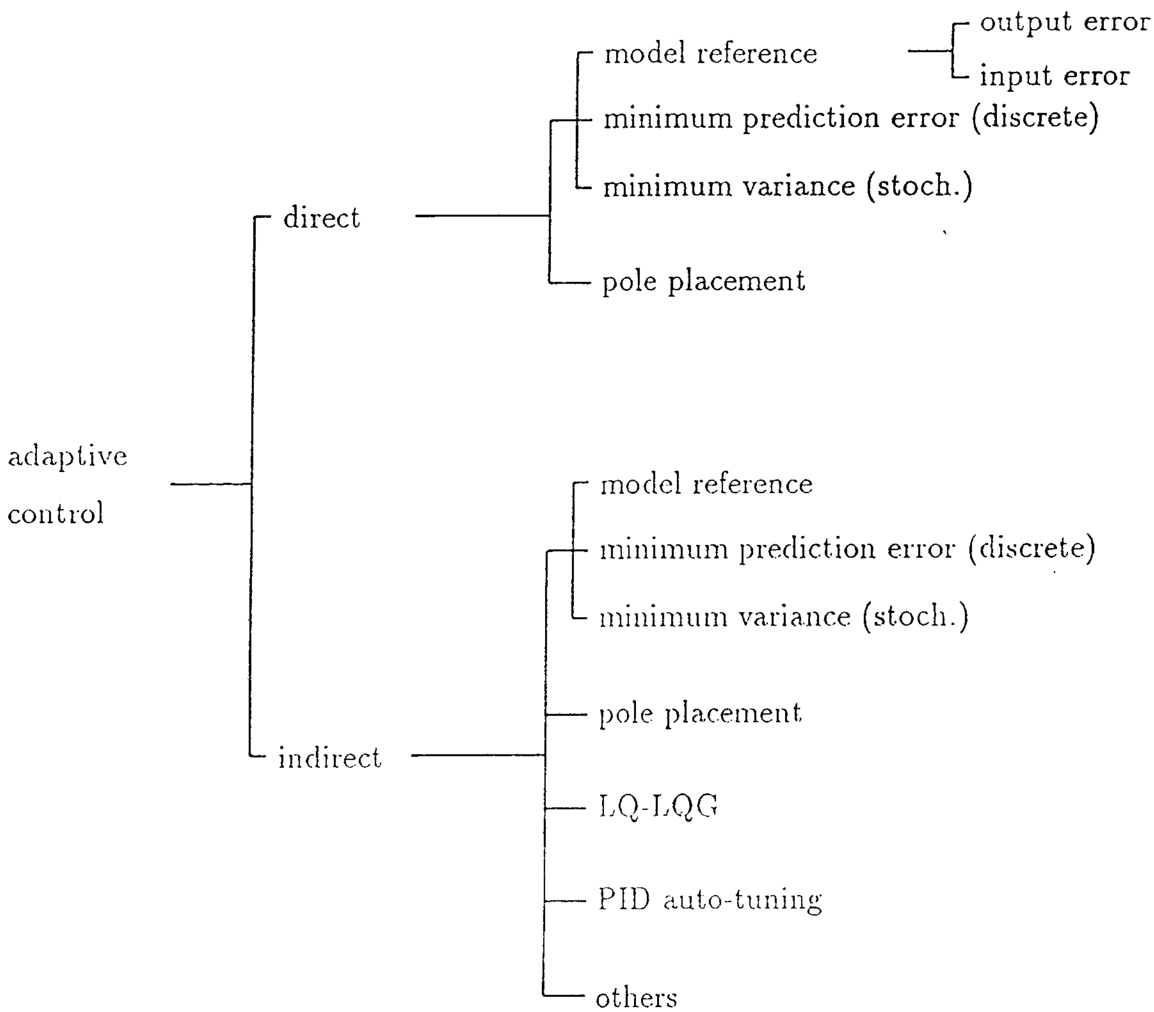

Figure 0.2: Adaptive controller design techniques

the reference model output. There are two different ways to implement a direct MRAC scheme: one uses the output error (the difference between the output of the plant and the output of the reference model) and assumes that the reference model is strictly positive real (SPR), (*cf.* Definition 1.5); the other one uses the so-called input error (the difference between the inverse of the reference model applied to the output of the plant and the reference input) and does not require such assumption. The main restriction with the MRAC schemes is the fact that they are applicable only to minimum phase systems (systems whose zeros are in the open left-half plane). Another popular approach used with deterministic discrete-time models is the *minimum prediction error adaptive control*, where the control law attempts to minimize a cost function based on the square of the error between the predicted output of the closed-loop system and the desired output. Often, a linear combination of squared future commands is added to the cost

function to reduce the amplitude of the transients and avoid saturation of the input. Discrete-time MRAC can be reformulated as a particular case of minimum prediction error adaptive control, where the desired output is generated by a reference model. Depending on the cost function, a minimum phase assumption on the system may or may not be required. When stochastic models are used instead of deterministic models, this type of technique is called *minimum variance adaptive control.* In the literature, minimum variance adaptive control is often called self-tuning control. Unfortunately, the term self-tuning control is also used to designate minimum prediction error adaptive control, indirect adaptive control, or sometimes any method different than MRAC. Finally, another popular method is *adaptive pole placement*, which can be applied to minimum phase and non-minimum phase systems.

Conceivably, indirect adaptive control allows the use of almost any type of control methods and any type of estimation technique. For example, it is possible to implement adaptive linear quadratic (LQ) and linear quadratic gaussian (LQG) regulators, or simply auto-tuning PID's. However, the considerable flexibility of the indirect adaptive control approach in the choice of both the controller and the identifier is greatly offset by the difficulty to guarantee stability. Indeed, for most of the indirect adaptive control schemes, either stability is not guaranteed, or the reference input must satisfy sufficient excitation constraints with the controller parameters updated only when adequate information has been obtained from the identifier. Estimates of the parameters must also be constrained to known (possibly disjoint) convex sets in which no unstable pole-zero cancellations occur. Unfortunately, to require sufficient excitation in the reference input is often not realistic. On the other hand, stability can be guaranteed more easily for (direct) model reference adaptive control schemes, minimum prediction error adaptive control schemes, and minimum variance adaptive control schemes (self-tuning regulators), with adequate *a priori* structural knowledge but without any particular constraints on the reference input signals.

Finally, note that there exist several relatively recent books studying in great details SISO adaptive control techniques, *e.g.*, the book by Goodwin & Sin [1] and the book by Åström & Wittenmark [2] about adaptive control of discrete-time deterministic and stochastic systems, or the book by Sastry & Bodson [3] and the book by Narendra & Annaswamy [4] about adaptive control of continuous-time systems.

0.2 Multivariable adaptive control

Survey of theoretical results

During the past ten years, adaptive control techniques for SISO systems were extended to multivariable systems. Several MIMO adaptive control algorithms based on the direct model reference approach have been proposed. Using continuous-time models, Monopoli & Hsing [5] first presented a direct MRAC scheme for a considerably restricted class of MIMO systems.

Then, Elliot & Wolovich [6] proposed a direct MRAC input error scheme for minimum phase MIMO systems, Das [7] a direct MRAC input error scheme requiring less *a priori* knowledge on the system, Tao & Ioannou [8] a robust input error scheme, and Singh & Narendra [9, 10] a direct MRAC output error scheme. A description of the input error scheme is also found in the book by Sastry & Bodson [3], and of the output error scheme in the book by Narendra & Annaswamy [4]. Discrete-time deterministic MRAC schemes were proposed by Elliot & Wolovich [11], Dugard *et al.* [12], Johansson [13], Ortega *et al.* [14], Das [7], and assuming that the state information is available, Kim & Bien [15]. Discrete-time stochastic MRAC schemes can be found in Bittanti & Scattolini [16], Scattolini & Clarke [17], and Scattolini [18].

In the deterministic discrete-time case, minimum prediction error adaptive controllers were also adapted to MIMO systems. Goodwin *et al.* [19] first proved global convergence of some prediction error adaptive control algorithms applied to a class of MIMO systems, then Goodwin & Long [20] extended the previous results to a more general class of MIMO systems. Finally, a modified version using a different *a priori* information was proposed by Tsiligiannis & Svoronos [21].

In the stochastic discrete-time case, MIMO minimum variance adaptive controllers were first proposed, by Keviczky & Hetthessy [22] without stability analysis, for restricted classes of systems, by Borison [23], Koivo [24], Goodwin *et al.* [25], Keviczky & Kumar [26], Bayoumi *et al.* [27], then for a general class of stochastic systems by Dugard *et al.* [28], Han *et al.* [29], Scattolini & Clarke [17], and Scattolini [18]. The MIMO minimum prediction error and minimum variance adaptive control algorithms are also described in the book by Goodwin & Sin [1].

The pole placement approach was investigated in the stochastic case by Prager & Wellstead [30] and Mikleš [31], and in the deterministic case by Elliot *et al.* [32], Djaferis *et al.* [33], Elliot & Wolovich [11], and Willner *et al.* [34].

Several auto-tuning regulators for MIMO systems have also been proposed, *e.g.*, by Davison [35], Penttinen & Koivo [36], Jones & Porter [37], and Miller & Davison [38]. Finally, we should also mention that a paper by Dugard & Dion [39] surveys direct adaptive control methods for MIMO systems.

Applications

There exist several applications of adaptive control to multivariable systems in process control. Most of these implementations have used the minimum variance adaptive controller, often called self-tuning regulator (STR). The earliest implementation is probably the application of a STR to a cement raw material blending process by Keviczky *et al.* [40]. A STR has also been used for the control of a binary distillation column by Morris *et al.* [41], and for the control of a drum boiler by Fessl [42]. Simulations with a STR applied to various processes have also been made and are described in papers previously cited: application of a STR to the head-box of a paper machine in Borison [23] and Koivo [24], to a hydraulic system in Prager & Wellstead [30],

to a binary distillation column in Bittanti & Scattolini [16], and to an aircraft in Kim & Bien [15]. A minimum prediction error adaptive controller has also been applied by Martin-Sanchez & Shah [43] to a binary distillation column.

There is also growing interest in the application of indirect adaptive control schemes to flight control systems, in particular to the problem of reconfiguration in the presence of failures, *e.g.*, Dittmar [44], Morse & Ossman [45], Gross & Migyanko [46], and Chandler *et al.* [47]. Large space structures control is another area where applications of multivariable adaptive methods are expected to grow, *e.g.*, Bayard *et al.* [48] and Bayard [49].

0.3 Limitations of current multivariable algorithms

0.3.1 Structural information

One of the main problems with the current parametric methods for MIMO adaptive control is the amount of required *a priori* knowledge about the structure of the system. A typical transfer function for a continuous-time deterministic SISO system will have the following form

$$p(s) = k_p \frac{b(s)}{a(s)} = k_p \frac{s^m + b_{m-1}s^{m-1} + \ldots + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_0} \tag{0.1}$$

where n is the *order* of the system, $n-m$ is the *relative degree* of $p(s)$, and k_p is the *high-frequency gain*. Similar definitions hold for discrete-time systems, except that the relative degree is called *time delay* and the high-frequency gain is called *time delay gain*. In SISO MRAC, minimum prediction error adaptive control, and minimum variance adaptive control, *a priori* knowledge consists in an upper bound on the order of system and the relative degree of the transfer function for continuous-time systems, or the time delay for discrete-time systems. Furthermore, knowledge of the sign of the high-frequency gain for continuous-time systems or the time delay gain for discrete-time systems is also often required to prove stability. For SISO adaptive pole placement algorithms and for most indirect adaptive control algorithms, the required *a priori* knowledge is the order of the system.

In MIMO adaptive control, the equivalent *a priori* knowledge is much more complex. The earliest methods dealt with the problem by restricting their applicability to some very particular class of multivariable systems. Later on, the first direct adaptive control (MRAC, minimum prediction error adaptive control algorithms, and minimum variance adaptive control) algorithms dealing with general classes of system were developed. They required the full knowledge of the *Hermite normal form* or the *interactor matrix* (see Definition 1.4), and also the knowledge of an upper bound on the maximum of the *observability indices* (see Definition 1.6). The knowledge of the interactor matrix may be seen as equivalent to the knowledge of the relative degree or the time delay in the SISO case. Indeed, the interactor matrix becomes simply $(s+a)^{n-m}$ (where a is an arbitrary positive constant) in the case of an SISO continuous-time

system (0.1). The maximum of the observability indices is equal to the order of the system in the SISO case. MIMO algorithms are found in the papers by Goodwin & Long [20], Elliot & Wolovich [6], Singh & Narendra [9, 10], Elliot & Wolovich [11], Bittanti & Scattolini [16], and Scattolini [18]. Furthermore, to guarantee stability these algorithms require the knowledge of the *high-frequency gain matrix* (see Definition 1.4) or the knowledge of a matrix whose product with the high-frequency gain matrix is positive definite (in the SISO case, this is equivalent to the knowledge of the sign of the high-frequency gain). Note that the discrete-time algorithm in Goodwin & Long [20] does not explicitly require the knowledge of the time delay gain matrix. However, this knowledge will be necessary for a practical implementation (see Chapter 4).

Additional complexity comes from the fact that the interactor matrix is not necessarily a diagonal matrix with known polynomials on the diagonal, but may contain off-diagonal coefficients that depend on the unknown parameters of the systems. Therefore, requiring the *a priori* knowledge of the interactor matrix often makes impractical the application of these algorithms. In recent work, the full knowledge of the interactor matrix has been progressively reduced to the knowledge of a small set of integer values: the relative degrees or delays on the diagonal of the interactor matrix (which can be seen as an extension of the relative degree or the time delay in the SISO case), see, *e.g.*, Elliot & Wolovich [11], Johansson [13], Dugard *et al.* [12], Dugard *et al.* [28], Scattolini & Clarke [17], Ortega *et al.* [14], Das [7], and Dion *et al.* [50]. However, stability is not proven for these algorithms, or it is implicitly assumed that the high-frequency gain matrix is at least partially known (see Chapter 5).

Turning to direct adaptive pole placement, we note that the algorithms proposed by Elliot *et al.* [32] and Willner *et al.* [34] require the *a priori* knowledge of the *controllability indices* (see Definition 1.7) and a bound on the maximum of the observability indices. Other direct adaptive pole placement algorithms, requiring less *a priori* knowledge, either are limited to a particular class of systems (*cf.* Prager & Wellstead [30]), or are not guaranteed to be stable (*cf.* Djaferis *et al.* [33]). Indirect adaptive pole placement algorithms, on the other hand, will require the knowledge of the observability indices and a upper bound on the maximum of the controllability indices (*cf.* Elliot & Wolovich [11]). Finally, indirect adaptive control algorithms generally require the knowledge of the observability indices.

It is not clear how the structural knowledge (interactor matrix, observability indices, controllability indices, high-frequency gain matrix) can be obtained if the system is unknown. The quantities are far less intuitive than SISO indices such as system order and time delay. Quantities like the high-frequency gain matrix and the non-diagonal elements of the Hermite normal form (or the interactor matrix) depend on the unknown parameters of the system. Structural indices like the observability indices or the controllability indices are not determined by more intuitive indices like system order, time delays or relative degrees of the elements of the transfer function matrix. It is often argued that these indices *generically* will satisfy certain constraints (*e.g.*, all observability indices are equal if the order is a multiple of the number of outputs). However, real systems often depart from genericity because of strong physical constraints, for example, a cross-coupling term in a transfer function matrix that is a true zero because one of

the inputs has no influence on one of the outputs. Current MIMO adaptive algorithms requiring less *a priori* information often do not guarantee stability. Therefore, the applicability of adaptive control to unknown multivariable systems is limited.

0.3.2 Parameter convergence and robustness

Even when stability is guaranteed, another problem with current MIMO direct adaptive control algorithms is that parameter convergence properties have not been established. In fact, parameter convergence *cannot* be guaranteed for most existing algorithms because non-unique parameterizations are used. In parametric adaptive control, the controller is parameterized so that a vector of controller parameters corresponds to each plant. If the controller parameters are not uniquely defined for each plant, the estimated parameters cannot converge to unique values. Even if the input signals are sufficiently rich (contain enough frequencies), the estimated parameters will not converge to a unique value but will converge to a set instead. The lack of parameter convergence properties is a potential problem for the robustness of these schemes to noise and unmodeled dynamics. Indeed, since the set to which the parameter converge is usually unbounded, small measurement noise and unmodeled dynamics can force convergence of the parameters to regions of the parameter space where the system becomes unstable. Conversely, if the controller parameters are uniquely defined and the inputs sufficiently rich, the parameters remain in the neighborhood of their nominal value and a certain degree of robustness is guaranteed (*cf.* Sastry & Bodson [3]).

In indirect adaptive control, the convergence issue has an equal importance for the robustness of the algorithms. Furthermore, for several indirect adaptive control schemes, *e.g.*, indirect adaptive pole placement, stability will not be guaranteed unless convergence can be assured.

0.4 Contributions

The unrealistic amount of *a priori* knowledge, the lack of stability proofs, and the lack of parameter convergence results in current multivariable adaptive control algorithms prompted the undertaking of the research presented in this thesis. The decision was made to focus on continuous-time systems since results obtained with continuous-time adaptive control schemes can often be transferred to the corresponding discrete-time schemes with little difficulty, while the reverse is not true.

1. *Parameter convergence*
   The first problem addressed in this thesis is that of parameter convergence. We define *identifiable* parameterizations for direct MRAC, direct adaptive pole placement (PPAC), and for a recursive identifier (that could be used in an indirect adaptive control scheme). The word identifiable is used here in analogy to the concept of *identifiability* in parameter estimation. A parameterization is a representation of a class of systems which associates

a vector of parameters to each system. For a linear time invariant system, one will usually says that a parameterization is identifiable if each transfer function corresponds to a unique parameter vector. Similarly, in direct adaptive control, the controller is parametrized so that a vector of controller parameters corresponds to each plant. Then, one may say that the parameterization is identifiable if the controller parameters are uniquely defined for each plant. Once an identifiable parameterization has been defined, then, assuming that the adaptive system is stable, parameter convergence can be guaranteed if persistency of excitation conditions are met. The implication of the results on robustness are illustrated with an example that shows an MRAC system with noise and unmodeled dynamics where a non-identifiable parameterization leads to instability and an identifiable parameterization does not.

2. *Simple frequency domain tests for parameter convergence*
   Another contribution of this work is the derivation of simple frequency domain conditions on the inputs that guarantee that the persistency of excitation conditions are satisfied and the parameters converge to their nominal value. An important difference with SISO systems is that the sufficient condition for parameter convergence may be different from the necessary condition. Indeed, parameter convergence may depend on the location of the spectral components of the inputs. The frequency domain conditions are derived for a direct MRAC scheme, a direct PPAC scheme, and a recursive identifier. Although the parameterizations under consideration in the three different schemes are all identifiable, they have different numbers of parameters and requirements for parameter convergence. Several examples are presented illustrating these differences and showing cases where a direct scheme converges and an indirect scheme does not, and *vice versa*. The so-called equivalence between direct and indirect schemes is found to be more complicated than it appears from the SISO case.

3. *Stable MRAC with unknown high-frequency gain matrix*
   Another important contribution of this work is the relaxation of the requirement of knowledge of the high-frequency gain matrix in MRAC. In most potential applications of multivariable adaptive control, the high-frequency gain matrix depends on the unknown parameters. Therefore, to require its knowledge is in most cases an unrealistic constraint. If we assume that the high-frequency gain matrix is unknown and is estimated, for stability reasons, its estimate must stay nonsingular. If the estimate approaches a singularity region, depending on the type of MRAC scheme, either adaptation may stop and there is a possibility of instability, or the control law may become undefined. Our result is achieved by applying a particular transformation to the estimated parameters to obtain the controller parameters. This transformation does not modify the properties of classic MRAC algorithms and allows us to prove stability without requiring knowledge of the high-frequency gain (only an upper bound or a lower bound on the norm of the high-frequency gain is required). Illustrative examples show that the parameter transfor-

mation is absolutely necessary if the high-frequency gain matrix is unknown, so that the singularity of the estimate of the high-frequency gain matrix is more than an academic issue. Indeed, the examples show MRAC systems whose output error converges with the parameter transformation but does not converge without it.

4. *Stable MRAC with unknown Hermite normal form or interactor matrix*
   The knowledge of the full Hermite form or the interactor matrix is also relaxed to the knowledge of the relative degrees of the diagonal elements of the Hermite normal form. This is achieved by identifying the off-diagonal elements of the Hermite normal form or the interactor matrix. Further, by using the parameter transformation, we are able to prove stability without requiring the knowledge of the high-frequency gain matrix nor the knowledge of the interactor matrix. This result considerably increases the applicability of MRAC to multivariable systems with respect to current algorithms.

5. *Observability and pseudo-observability indices*
   Finally, we show that for direct MRAC, direct PPAC, as well as indirect schemes, the knowledge of the observability indices can be replaced by the knowledge of any set of pseudo-observability indices (see Definition 1.6), which is a less restrictive requirement. Simple criteria for the identification of the observability indices and the pseudo-observability indices are also presented in the thesis. These results were obtained by working with canonical and pseudo-canonical matrix fraction descriptions which are uniquely defined for a given system. The criteria for structural identification (identification of the canonical and pseudo-canonical observability indices) use information from the recursive identifier to decide if the correct structural indices have been selected. Therefore, on-line structure selection becomes possible. The advantage of using pseudo-observability indices is the fact that they define overlapping classes of systems as opposed to the observability indices which define distinct classes of systems. Therefore, the pseudo-observability indices can change without having to re-identify completely the parameters (by a similarity transformation, the estimates of the parameters under the old structural indices can be transformed into estimates under the new structural indices).

0.5 Organization of the thesis

The thesis is divided into five chapters. A brief outline of the content of the chapters is as follows. In the first chapter, concepts from multivariable system theory are recalled. Furthermore, lesser known facts about canonical and pseudo-canonical forms are emphasized. The second chapter presents current MIMO adaptive control algorithms and their properties. A direct MRAC scheme, a direct PPAC scheme, and a recursive identifier are considered. In the third chapter, it is shown how to define identifiable parameterizations for the three schemes under consideration in the previous chapter (MRAC, PPAC, and recursive identification). The third

chapter also gives the frequency domain conditions for parameter convergence for these three schemes. The parameter transformation relaxing the MRAC requirement of the knowledge of the high-frequency gain matrix is developed in the fourth chapter. The chapter provides also a complete proof of stability of the MRAC scheme with the parameter transformation. Structure identification is presented in the fifth chapter. It is shown how to introduce the identification of the off-diagonal elements of the interactor matrix in a MRAC control scheme. A stability proof is given without requiring knowledge of the high-frequency gain matrix. Finally, criteria for on-line validation and selection of the pseudo-observability indices and the observability indices are also developed in the fifth chapter.

Chapter 1

Multivariable system theory

1.0 Introduction

This chapter introduces the mathematical notation used in this work as well as useful properties and definitions. It concentrates on results and definitions related to multivariable systems theory and, more particularly, multivariable system descriptions. There exist three different types of multivariable system descriptions. Each of them makes up a section of this chapter. The first is the transfer function matrix description. This is the direct transposition of the notion of transfer function in SISO systems. The very important concepts of interactor matrix and Hermite normal form used in MIMO MRAC control are introduced in this section. The second section discusses MIMO state-space descriptions. This section also recalls the definition of the observability indices and the controllability indices and the lesser known pseudo-observability indices and pseudo-controllability indices. These structural indices, which are unique to MIMO systems, are very important for the structure of the controllers used in MIMO parametric adaptive control. Further, the section gives some canonical forms that will be used later on to define identifiable controllers. Finally, the last section describes matrix fraction descriptions which allow us to manipulate matrices of polynomials instead of matrices of rational functions. Canonical matrix fraction descriptions are also described in the section. Furthermore, it is shown how they relate to canonical state-space forms. Generally, more details may be found in the book by Kailath [51] on these topics, except for pseudo-canonical forms which are described in Correa & Glover [52, 53] and Gevers & Wertz [54]. Multivariable parameterizations and canonical forms are also presented in greater details in de Mathelin [55]. For the reader who is not very familiar with multivariable system theory, we tried to alleviate the intrinsic dryness of this chapter by introducing numerous examples.

1.1 Notation

In this work, $P(s)$ denotes the transfer function of a linear time invariant operator, P. Then, $P[u]$ denotes the output of the operator in the time domain with input $u(t)$. Furthermore, the dependency in s of a transfer function or the dependency in t of a signal will be often omitted when there is no ambiguity. Finally, $\Re^{p\times m}(s)$ will be the set of $(p \times m)$ matrices whose elements are rational functions of s and $\Re^{p\times m}[s]$ will be the set of $(p \times m)$ matrices whose elements are polynomial functions of s.

1.2 Transfer function matrix description

The transfer function matrix is the direct transposition of the notion of transfer function in the single input single output (SISO) case to the multi-input multi-output (MIMO) case. An m input p output, linear, time-invariant system can be described by

$$y_p(s) = P(s)u(s) \tag{1.1}$$

where $P(s) \in \Re^{p\times m}(s)$, $y_p(s)$ is the Laplace transform of the output of the system, and $u(s)$ the Laplace transform of the input. $P(s)$ is the *transfer function matrix* of the system and equation (1.1) is called a *transfer function matrix description* of the system.

Definition 1.1 : Proper and strictly proper transfer function matrix
A transfer function matrix $P(s)$ is proper iff:

$$\lim_{s\to\infty} P(s) \ \text{ exists}$$

and is strictly proper iff:

$$\lim_{s\to\infty} P(s) \ = \ 0$$

Definition 1.2 : Markov parameters
Let $P(s) \in \Re^{p\times m}$ be strictly proper. The $(p \times m)$ scalar matrices $\{h^{(i)}\}$ such that:

$$P(s) = \sum_{i=1}^{\infty} h^{(i)} s^{-i} \tag{1.2}$$

are called the Markov parameters of the system.

Based on the above definition, the Markov parameters can be computed the following way:

$$\begin{aligned} h^{(0)} &= 0 \\ h^{(1)} &= \lim_{s\to\infty} \ sP(s) \end{aligned}$$

$$
\begin{aligned}
h^{(2)} &= \lim_{s\to\infty} \; s[sP(s) - h^{(1)}] \\
h^{(3)} &= \lim_{s\to\infty} \; s[s^2P(s) - sh^{(1)} - h^{(2)}] \\
&\vdots \\
h^{(i)} &= \lim_{s\to\infty} \; [s^iP(s) - \sum_{j=0}^{i-1} s^{i-j}h^{(j)}]
\end{aligned}
$$

If we define $P(t)$ as the impulse response of the system, $P(t) = \mathcal{L}^{-1}\{P(s)\}$, then by the properties of the Laplace transform, the Markov parameters are equal to the value at the origin of $P(t)$ and its successive derivatives:

$$
h^{(i)} = \frac{d^{i-1}}{dt^{i-1}} P(t)_{|t=0}
$$

The Markov parameters define a system uniquely, in the sense that for one set of Markov parameters there exists only one system. Also, the Markov parameters are invariants of the system, *i.e.*, one particular system has only one set of Markov parameters.
Example: the 2×2 transfer function matrix

$$
P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}
$$

is strictly proper and has the following first four Markov parameters

$$
h^{(0)} = 0 \qquad h^{(1)} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \qquad h^{(2)} = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \qquad h^{(3)} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}
$$

Definition 1.3 : Hankel matrices and order of the system

The matrices:

$$
H_{[i,j]} = \begin{bmatrix} h^{(i)} & h^{(i+1)} & \dots & h^{(i+j-1)} \\ h^{(i+1)} & h^{(i+2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ h^{(i+j-1)} & \dots & \dots & h^{(i+2j-2)} \end{bmatrix} \tag{1.3}
$$

where the $\{h^{(i)}\}$ are the Markov parameters, are called the Hankel matrices.

The infinity Hankel matrix is the Hankel matrix $H_{[0,\infty]}$ and, like the Markov parameters, uniquely defines the system. Finally, the order, n, of the system is the rank of the infinity Hankel matrix.

Definition 1.4 : Hermite normal form, interactor matrix, and high-frequency gain matrix

Morse [56] showed that for any square, strictly proper, and nonsingular plant $P(s) \in \Re^{p \times p}(s)$, there exists a *unique* matrix $H(s) \in \Re^{p \times p}(s)$, called the *Hermite normal form*, such that:

$$\begin{aligned} P(s) &= H(s)U(s) \\ U(s) &\in \Re^{p \times p}(s) \text{ and } \lim_{s \to \infty} U(s) = K_p \quad \text{nonsingular} \\ H(s) &= \begin{bmatrix} \frac{1}{(s+a)^{r_1}} & 0 & \cdot & \cdot \\ \frac{h_{21}(s)}{(s+a)^{r_2-1}} & \frac{1}{(s+a)^{r_2}} & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \frac{1}{(s+a)^{r_p}} \end{bmatrix} \end{aligned}$$

where $\partial h_{ij}(s) < r_i - 1$ and $a > 0$ is arbitrary, but fixed *a priori*. The matrix K_p is the *high-frequency gain* matrix. The integers r_i extend the notion of relative degree r of an SISO transfer function.

It can also be shown, see Sastry & Bodson [3], that the inverse of the Hermite normal form is the *interactor matrix* $\xi(s)$ used in several model reference adaptive control algorithms (*e.g.*, Elliot & Wolovich [32] and defined in Wolovich & Falb [57] as the $(p \times p)$ polynomial matrix such that:

$$\begin{aligned} \lim_{s \to \infty} \xi(s)P(s) &= K_p \quad \text{nonsingular} \\ \xi(s) &= (\Sigma^*(s)F(s) + I)\Delta(s) \end{aligned}$$

with

$$\begin{aligned} \Delta(s) &= \text{diag}\{(s+a)^{r_i}\} \\ F(s) &= \text{diag}\{(s+a)\} \end{aligned} \quad \text{and} \quad \Sigma^*(s) = \begin{bmatrix} 0 & 0 & \cdot & \cdot \\ \sigma_{21}(s) & 0 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{p1}(s) & \cdot & \sigma_{pp-1}(s) & 0 \end{bmatrix}$$

It was also shown by Das [7] that the elements of $\Sigma^*(s)$ are zero or satisfy the condition

$$\partial \sigma_{ij} \leq \Big(\sum_{k=j+1}^{i} (r_k - 1) \Big) - 1 \qquad i > j \tag{1.4}$$

Finally, it should be noted that Dion *et al.* [50] proved that the bounds given in (1.4) are the best possible ones. In the sense that there exist systems such that the inequalities in (1.4) become equalities.

Example: given the following 2×2 transfer function matrix

$$P(s) = \begin{bmatrix} \frac{s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

then

$$
\begin{aligned}
H(s) &= \begin{bmatrix} \frac{1}{(s+a)} & 0 \\ \frac{-0.5}{(s+a)} & \frac{1}{(s+a)^2} \end{bmatrix} \\
\xi(s) &= H^{-1}(s) = \begin{bmatrix} (s+a) & 0 \\ 0.5(s+a)^2 & (s+a)^2 \end{bmatrix} \\
&= (\Sigma^*(s)F(s)+I)\Delta(s) = (\begin{bmatrix} 0 & 0 \\ 0.5 & 0 \end{bmatrix}(s+a)+I)\begin{bmatrix} (s+a) & 0 \\ 0 & (s+a)^2 \end{bmatrix} \\
K_p &= \begin{bmatrix} 0 & -2 \\ 0.5 & -0.5 \end{bmatrix}
\end{aligned}
$$

Therefore, $r_1 = 1$, $r_2 = 2$, and it can be checked that $\partial\sigma_{21} = (r_2 - 1) - 1 = 0$.

Definition 1.5 : Positive real transfer matrices
A $p \times p$ matrix $P(s)$ of functions of a complex variable $s = \sigma + j\omega$ is positive real (PR) if:

- $P(s)$ has elements that are analytic $\forall\ \sigma > 0$,
- $P^*(s) = P(s^*)$, $\forall\ \sigma > 0$,
- $P^T(s^*) + P(s) \geq 0$, $\forall\ \sigma > 0$.

Furthermore, $P(s)$ is strictly positive real (SPR) if $P(s-\epsilon)$ is PR for some $\epsilon > 0$.

Comment: For a rational transfer function with real elements, the main condition to be SPR is that $P(s)$ must be stable and $P^{*T}(j\omega) + P(j\omega) > 0$. SPR transfer functions form a rather restricted class of transfer functions.

1.3 State-space description

Given a m input p output strictly proper system of order n (*cf.* Definition 1.3), it can be shown that a *minimal* state-space realization $\{A, B, C\}$ has the form

$$
\begin{aligned}
\dot{x} &= Ax + Bu \\
y_p &= Cx
\end{aligned}
\tag{1.5}
$$

where $x(t)$ is a n dimensional state vector, $A \in \Re^{n\times n}$, $B \in \Re^{n\times m}$, and $C \in \Re^{p\times n}$. The state-space realization is related to the transfer function matrix description by

$$
P(s) = C(sI - A)^{-1}B \tag{1.6}
$$

There exists an infinity of such realizations, related by nonsingular state transformations (similarity transformations). Indeed, assuming that T is a $(n \times n)$ nonsingular matrix, then

$$A^* = T^{-1}AT \qquad B^* = T^{-1}B \qquad C^* = CT$$

define another minimal state-space realization of the system with state vector $x^* = T^{-1}x$.

Definition 1.6 : Observability and pseudo-observability indices

Suppose that $\mathcal{O}[C, A]$ is the $(pn \times n)$ observability matrix associated with the minimal realization $\{A, B, C\}$ without redundant outputs (*i.e.*, C has full rank). Then

$$\mathcal{O}[C, A] = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

Since the representation is minimal, it is completely observable and the matrix $\mathcal{O}[C, A]$ has full rank n. Therefore, one can select n linearly independent rows of $\mathcal{O}[C, A]$ to form a row basis for $\mathcal{O}[C, A]$. There exist many ways to select such a row basis. However, we can limit this number by first imposing that

1. Rows $1, \ldots, p$ belong to the basis (always possible, since it is assumed that there are no redundant outputs).

2. Row $(i-p)$, $i > p$, belongs to the basis if row i does so (also always possible).

Given these two constraints, there still exist several potential row bases. For example, if $p = 2$ and $n = 5$, then rows $\{1,2,3,4,5\}$, $\{1,2,3,4,6\}$, $\{1,2,3,5,7\}$, $\{1,2,4,6,8\}$ are potential row basis. For a particular system, not all these combinations of rows are necessarily linearly independent. However, there always exists at least one of those combinations of rows which forms a basis since the system is completely observable. The row basis can be characterized by the indices $\{\rho_i\}$ defined as the total number of rows in the basis associated with the output y_{p_i}. For example, if $p = 2$ and $n = 5$, then the row basis made of rows $\{1,2,3,4,5\}$ (3 rows associated with output y_{p_1} and 2 rows associated with output y_{p_2}) can also be characterized by the indices $\{3,2\}$. Obviously, based on the definition of the indices $\{\rho_i\}$, $1 \leq \rho_i \leq (n-p+1)$ and $\sum_{i=1}^{p} \rho_i = n$. The numbers $\{\rho_i\}$ are called the *pseudo-observability indices*. There generally exist several choices of pseudo-observability indices for a particular system. For example, if $p = 2$ and $n = 5$, the pseudo-observability indices could be one or several of the sets $\{3,2\}$, $\{2,3\}$, $\{4,1\}$, and $\{1,4\}$.

Now, a unique row basis can be defined by requiring that

3. The n first linearly independent rows must be selected.

Then, the corresponding set of indices, denoted $\{\nu_i\}$, are simply called *observability indices* and are uniquely defined. Clearly, one of the sets of pseudo-observability indices is the set of observability indices.

Finally, it can be shown (see, *e.g.*, Kailath [51], Chapter 6), that the sets of pseudo-observability indices and the observability indices are invariants of the system. In other words, they are independent of a particular state-space realization.

Example: given the system with transfer function matrix

$$P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

a minimal state-space realization is

$$A = \begin{bmatrix} -7 & -5 & 3 & 4 \\ 3 & 3 & -2 & 3 \\ -17 & -13 & 6 & -10 \\ -7 & -7 & 3 & -6 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 \\ 1 & -3 \\ -2 & -1 \\ -3 & 3 \end{bmatrix} \qquad C = \begin{bmatrix} 2 & 2 & -1 & 1 \\ 2 & 1 & -1 & 1 \end{bmatrix}$$

Therefore, the corresponding observability matrix

$$\mathcal{O}[C, A] = \begin{bmatrix} 2 & 2 & -1 & 1 \\ 2 & 1 & -1 & 1 \\ 2 & 2 & -1 & 2 \\ -1 & -1 & 1 & -1 \\ -5 & -5 & 2 & -4 \\ -6 & -4 & 2 & -3 \\ 14 & 12 & -5 & 9 \\ 17 & 13 & -7 & 10 \end{bmatrix}$$

It can be verified that rows $\{1, 2, 3, 4\}$ are linearly independent, as well as rows $\{1, 2, 3, 5\}$ and rows $\{1, 2, 4, 6\}$. Therefore, $\{2, 2\}$, $\{3, 1\}$, and $\{1, 3\}$ are the pseudo-observability indices of the system. $\{2, 2\}$ are the observability indices.

Definition 1.7 : Controllability and pseudo-controllability indices

Suppose that $\mathcal{C}[A, B]$ is the $(n \times mn)$ controllability matrix associated with the minimal realization $\{A, B, C\}$ without redundant inputs (*i.e.*, B has full rank). Then

$$\mathcal{C}[A, B] = \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}$$

Since the representation is minimal, it is completely controllable and the matrix $\mathcal{C}[A, B]$ has full rank n. Therefore, one can select n linearly independent columns of $\mathcal{C}[A, B]$ to form a

column basis for $C[A,B]$. By selecting the columns of the controllability matrix according to similar rules than when selecting the rows of the observability matrix, one can define sets of *pseudo-observability indices*, $\{\kappa_i\}$ and a unique set of *controllability indices*, $\{\mu_i\}$. Again, the sets of pseudo-controllability indices and the controllability indices are invariants of the system.

Example: given the system with the minimal state-space realization from Definition 1.6, the corresponding controllability matrix

$$C[A,B] = \begin{bmatrix} 0 & 0 & 1 & 0 & -2 & -1 & 3 & 3 \\ 1 & -3 & -2 & 2 & 2 & 0 & -2 & -2 \\ -2 & -1 & 5 & 3 & -11 & -8 & 22 & 19 \\ -3 & 3 & 5 & 0 & -8 & -5 & 15 & 13 \end{bmatrix}$$

It can be verified that columns $\{1,2,3,5\}$ are linearly independent, but columns $\{1,2,3,4\}$ and columns $\{1,2,4,6\}$ are not ($AB_2 + B_1 + B_2 = 0$ where B_i is the i-th column of B). Therefore, $\{3,1\}$ are the only pseudo-controllability indices of the system, and are also the controllability indices.

Definition 1.8 : Observability and pseudo-observability canonical forms

The state-space realization $\{A_{poy}, B_{poy}, C_{poy}\}$ obtained by applying to any minimal realization $\{A,B,C\}$, the state transformation $T = Q^{-1}[C,A]$, where $Q[C,A]$ is defined by

$$Q[C,A] = \begin{bmatrix} C_1 \\ C_1 A \\ \vdots \\ C_1 A^{\rho_1 - 1} \\ C_2 \\ \vdots \\ C_p A^{\rho_p - 1} \end{bmatrix}$$

where C_i is the i-th row of C, is called a *pseudo-observability canonical form*. If the observability indices are known and the $\{\rho_i\}$ are replaced by the $\{\nu_i\}$ in $Q[C,A]$, then the resulting realization $\{A_{oy}, B_{oy}, C_{oy}\}$ is the *observability canonical form*. Clearly, one of the pseudo-observability canonical form is identical to the observability canonical form. It can be shown (see, *e.g.*, Correa & Glover [52, 53]) that the pseudo-observability canonical form has the following structure

$$C_{poy} = \text{block diag}\left\{ \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}, (1 \times \rho_i), i = 1, \ldots, p \right\}$$

$$A_{poy} = \text{block diag}\left\{ \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}, (\rho_i \times \rho_i), i = 1, \ldots, p \right\}$$

$$
+\text{block diag}\left\{ \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, (\rho_i \times 1), i = 1, \ldots, p \right\} \begin{bmatrix} \alpha_{110} & \cdots & \alpha_{11(\rho_1-1)} & \alpha_{120} & \cdots & \alpha_{1p(\rho_p-1)} \\ \alpha_{210} & \cdots & \alpha_{21(\rho_1-1)} & \alpha_{220} & \cdots & \alpha_{2p(\rho_p-1)} \\ \vdots & & & & & \vdots \\ \alpha_{p10} & \cdots & \alpha_{p1(\rho_1-1)} & \alpha_{p20} & \cdots & \alpha_{pp(\rho_p-1)} \end{bmatrix}
$$

$$
B_{poy} = \begin{bmatrix} h_{11}^{(1)} & \cdots & h_{1m}^{(1)} \\ \vdots & & \vdots \\ h_{11}^{(\rho_1)} & \cdots & h_{1m}^{(\rho_1)} \\ h_{21}^{(1)} & \cdots & h_{2m}^{(1)} \\ \vdots & & \vdots \\ h_{p1}^{(\rho_p)} & \cdots & h_{pm}^{(\rho_p)} \end{bmatrix}
$$

where the matrices $h^{(i)}$ are the Markov parameters of the system and the coefficients α_{ijk} are such that

$$
C_i A^{\rho_i} = \sum_{j=1}^{p} \sum_{k=0}^{\rho_j - 1} \alpha_{ijk} C_j A^k
$$

for all minimal state-space realizations $\{A, B, C\}$. Furthermore, it can be shown (see, *e.g.*, Correa & Glover [52, 53]) that the coefficients α_{ijk} are invariants for a given set of pseudo-observability indices. Consequently, for a given system there exists a fixed number of sets of pseudo-observability indices and for each of these sets there exists only one pseudo-observability canonical form.

The observability canonical form is identical to the pseudo-observability canonical form if the $\{\rho_i\}$ are replaced by $\{\nu_i\}$. However, knowing that the observability indices are $\{\nu_i\}$ implies that

$$
\alpha_{ijk} = 0 \quad \text{if} \quad \begin{cases} k \geq \min(\nu_i, \nu_j) & \text{and} \;\; j \geq i \\ k \geq \min(\nu_i + 1, \nu_j) & \text{and} \;\; j < i \end{cases}
$$

For a given system there will exist only one observability canonical form.
For example, if $m = 2$, $p = 3$, $n = 6$, $\{\rho_i\} = \{3, 1, 2\}$, then

$$
C_{poy} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}
$$

$$
A_{poy} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \alpha_{110} & \alpha_{111} & \alpha_{112} & \alpha_{120} & \alpha_{130} & \alpha_{131} \\ \alpha_{210} & \alpha_{211} & \alpha_{212} & \alpha_{220} & \alpha_{230} & \alpha_{231} \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \alpha_{310} & \alpha_{311} & \alpha_{312} & \alpha_{320} & \alpha_{330} & \alpha_{331} \end{bmatrix} \qquad B_{poy} = \begin{bmatrix} h_{11}^{(1)} & h_{12}^{(1)} \\ h_{11}^{(2)} & h_{12}^{(2)} \\ h_{11}^{(3)} & h_{12}^{(3)} \\ h_{21}^{(1)} & h_{22}^{(1)} \\ h_{31}^{(1)} & h_{32}^{(1)} \\ h_{31}^{(2)} & h_{32}^{(2)} \end{bmatrix}
$$

and if $\{\nu_i\} = \{3, 1, 2\}$, then

$$C_{oy} = C_{poy} \qquad A_{oy} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \alpha_{110} & \alpha_{111} & \alpha_{112} & \alpha_{120} & \alpha_{130} & \alpha_{131} \\ \alpha_{210} & \alpha_{211} & 0 & \alpha_{220} & \alpha_{230} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \alpha_{310} & \alpha_{311} & \alpha_{312} & \alpha_{320} & \alpha_{330} & \alpha_{331} \end{bmatrix}. \qquad B_{oy} = B_{poy}$$

In the case of the system given as example in Definition 1.6

$$P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

the pseudo-observability indices are $\{2, 2\}$, $\{3, 1\}$, and $\{1, 3\}$, and the observability indices $\{2, 2\}$. The observability canonical form (which is equal to the pseudo-observability canonical form with indices $\{2, 2\}$) is given by

$$A_{oy} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & -2 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & -2 & -2 \end{bmatrix} \qquad B_{oy} = \begin{bmatrix} 1 & -2 \\ -2 & 1 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \qquad C_{oy} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The pseudo-observability canonical form with indices $\{3, 1\}$ is

$$A_{poy} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -4 & -4 & 2 \\ -1 & -2 & -1 & 0 \end{bmatrix} \qquad B_{poy} = \begin{bmatrix} 1 & -2 \\ -2 & 1 \\ 3 & 1 \\ 0 & 1 \end{bmatrix} \qquad C_{poy} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The pseudo-observability canonical form with indices $\{1, 3\}$ is

$$A_{poy} = \begin{bmatrix} -1 & -2 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -2 & -3 & -3 \end{bmatrix} \qquad B_{poy} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \qquad C_{poy} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Definition 1.9 : Controllability and pseudo-controllability canonical forms

The state-space realization $\{A_{pcy}, B_{pcy}, C_{pcy}\}$ obtained by applying to any minimal realization $\{A, B, C\}$, the similarity transformation $T = R^{-1}[A, B]$, where $R[A, B]$ is defined by

$$R[A, B] = \begin{bmatrix} B_1 & AB_1 & \ldots & A^{\kappa_1 - 1}B_1 & B_2 & \ldots & A^{\kappa_m - 1}B_m \end{bmatrix}$$

where B_i is the i-th column of B, is called a *pseudo-controllability canonical form*. If the controllability indices are known and the $\{\kappa_i\}$ are replaced by the $\{\mu_i\}$ in $R[A, B]$, then the resulting realization $\{A_{cy}, B_{cy}, C_{cy}\}$ is the *controllability canonical form*. It can be shown that the pseudo-controllability canonical form has the following structure

$$
\begin{aligned}
B_{pcy} &= \text{block diag}\left\{ \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}, (\kappa_i \times 1), i = 1, \ldots, m \right\} \\
A_{pcy} &= \text{block diag}\left\{ \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, (\kappa_i \times \kappa_i), i = 1, \ldots, m \right\} \\
&+ \begin{bmatrix} \beta_{110} & \cdots & \beta_{m10} \\ \vdots & & \vdots \\ \beta_{11(\kappa_1-1)} & & \beta_{m1(\kappa_1-1)} \\ \beta_{120} & & \beta_{m20} \\ \vdots & & \vdots \\ \beta_{1m(\kappa_m-1)} & \cdots & \beta_{mm(\kappa_m-1)} \end{bmatrix} \text{block diag}\left\{ \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}, (1 \times \kappa_i), i = 1, \ldots, m \right\} \\
C_{pcy} &= \begin{bmatrix} h_{11}^{(1)} & \cdots & h_{11}^{(\kappa_1)} & h_{12}^{(1)} & \cdots & h_{1m}^{(\kappa_m)} \\ \vdots & & \vdots & \vdots & & \vdots \\ h_{p1}^{(1)} & \cdots & h_{p1}^{(\kappa_1)} & h_{p2}^{(1)} & \cdots & h_{pm}^{(\kappa_m)} \end{bmatrix}
\end{aligned}
$$

where the matrices $h^{(i)}$ are the Markov parameters of the system and the coefficients β_{ijk} are such that

$$A^{\kappa_i} B_i = \sum_{j=1}^{m} \sum_{k=0}^{\kappa_j - 1} \beta_{ijk} A^k B_j$$

for all minimal state-space realizations $\{A, B, C\}$. Furthermore, the coefficients β_{ijk} are invariants for a given set of pseudo-controllability indices.

The controllability canonical form is identical to the pseudo-controllability canonical form if the $\{\kappa_i\}$ are replaced by $\{\mu_i\}$. However, in the case of the canonical form, it can also be deduced that

$$\beta_{ijk} = 0 \quad \text{if} \quad \begin{cases} k \geq \min(\mu_i, \mu_j) \text{ and } j \geq i \\ k \geq \min(\mu_i + 1, \mu_j) \text{ and } j < i \end{cases}$$

In the case of the system given as example in Definition 1.6

$$P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

the pseudo-controllability and controllability indices are $\{3, 1\}$. The controllability canonical form (which is also the pseudo-controllability canonical form with indices $\{3, 1\}$) is given by

$$A_{cy} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix} \qquad B_{cy} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad C_{cy} = \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

1.4 Matrix fraction description

Given an m input p output system of order n with transfer function matrix $P(s) \in \Re^{p\times m}(s)$, a *matrix fraction description* (MFD) of the system is

$$y_p(s) = D_L(s)^{-1}N_L(s)u(s) = N_R(s)D_R^{-1}(s)u(s) = P(s)u(s) \tag{1.7}$$

where $D_L(s) \in \Re^{p\times p}[s]$, $N_L(s) \in \Re^{p\times m}[s]$, $D_R(s) \in \Re^{m\times m}[s]$, and $N_R(s) \in \Re^{p\times m}[s]$. Furthermore, $\{N_L(s), D_L(s)\}$ is called a left matrix fraction description (MFD) of the system, and $\{N_R(s), D_R(s)\}$ a right MFD. MFD's are not unique. Indeed, if we multiply on the left (right) a left (right) MFD by a nonsingular polynomial matrix, we get another left (right) MFD.

Definition 1.10 : Row and column degrees, row and column reduced matrix
Let $D(s) \in \Re^{k\times k}[s]$. Let $\partial r_i D$ denote the maximum polynomial degree in the i-th row of $D(s)$, and $\partial c_i D$ the maximum polynomial degree in the i-th column of $D(s)$. Then, $\exists$ two unique real $(k \times k)$ matrices $\Gamma_r[D]$ and $\Gamma_c[D]$, such that $\forall a \in \Re$, (arbitrary)

$$D(s) = \begin{cases} S_r(s)\Gamma_r[D] + D_r(s) \\ \Gamma_c[D]S_c(s) + D_c(s) \end{cases}$$

$$S_r(s) = \mathrm{diag}\left\{\ (s+a)^{\partial r_i D}\ \right\} \qquad S_c(s) = \mathrm{diag}\left\{\ (s+a)^{\partial c_i D}\ \right\}$$

where $D_r(s)$ and $D_c(s)$ are polynomial matrices with lower degree terms such that:

$$\begin{aligned} \partial r_i D_r &< \partial r_i D \\ \partial c_i D_c &< \partial c_i D \end{aligned}$$

The coefficients $\partial r_i D$ are called the *row degrees* of $D(s)$ and the coefficients $\partial c_i D$ the *column degrees* of $D(s)$. The matrix $\Gamma_r[D]$ is called the *leading row coefficient matrix* of $D(s)$ and $\Gamma_c[D]$

the *leading column coefficient matrix* of $D(s)$.
This result is illustrated by the following example:

$$\begin{aligned}
D(s) = \begin{bmatrix} 2s^2 - s + 3 & 3s - 2 \\ s^2 + 1 & 2 \end{bmatrix} &= \begin{bmatrix} 2s^2 + 4s + 2 & 0 \\ s^2 + 2s + 1 & 0 \end{bmatrix} + \begin{bmatrix} -5s + 1 & 3s - 2 \\ -2s & 2 \end{bmatrix} \\
&= \begin{bmatrix} (s+1)^2 & 0 \\ 0 & (s+1)^2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} -5s + 1 & 3s - 2 \\ -2s & 2 \end{bmatrix}
\end{aligned}$$

$$\Rightarrow \quad \Gamma_r[D] = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} 2s^2 + 4s + 2 & 3s + 3 \\ s^2 + 2s + 1 & 0 \end{bmatrix} + \begin{bmatrix} -5s + 1 & -5 \\ -2s & 2 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} (s+1)^2 & 0 \\ 0 & (s+1) \end{bmatrix} + \begin{bmatrix} -5s + 1 & -5 \\ -2s & 2 \end{bmatrix}
\end{aligned}$$

$$\Rightarrow \quad \Gamma_c[D] = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$$

A polynomial matrix $D(s)$ is said to be *row (column) reduced* if its leading coefficient matrix $\Gamma_r[D]$ ($\Gamma_c[D]$) is nonsingular. For example,

$$D(s) = \begin{bmatrix} 2s^2 - s + 3 & 3s - 2 \\ s^2 + 1 & 2 \end{bmatrix} \quad \Rightarrow \quad \Gamma_r[D] = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}$$

$$\Rightarrow \quad \Gamma_c[D] = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$$

is column reduced but not row reduced.
It can also be shown that, if $D(s)$ is row (column) reduced then

$$\deg \det\{D(s)\} = \sum_i \partial r_i D \qquad (\sum_i \partial c_i D)$$

and if $D(s)$ is not row (column) reduced

$$\deg \det\{D(s)\} < \sum_i \partial r_i D \qquad (\sum_i \partial c_i D)$$

Property 1.1 : **Strictly proper MFD**
Let $\{N_L(s), D_L(s)\}$ be a left MFD of $P(s)$, and $\{N_R(s), D_R(s)\}$ a right MFD, then

$$P(s) \quad \text{is strictly proper} \quad \Rightarrow \quad \left\{ \begin{array}{l} \partial r_i N_L < \partial r_i D_L \ \ \forall i \\ \partial c_i N_R < \partial c_i D_R \ \ \forall i \end{array} \right.$$

Furthermore, if $D_L(s)$ is row reduced and $D_R(s)$ column reduced, then

$$P(s) \quad \text{is strictly proper} \quad \Leftrightarrow \quad \begin{cases} \partial r_i N_L < \partial r_i D_L & \forall i \\ \partial c_i N_R < \partial c_i D_R & \forall i \end{cases}$$

Definition 1.11 : Coprime MFD

A left (right) MFD is said to be *coprime* if its greatest common left (right) divisors are *unimodular* matrices. A polynomial matrix is called unimodular, if its inverse is also a polynomial matrix, *i.e.*, if its determinant is a scalar different from zero. A left (right) MFD can always be transformed in a coprime left (right) MFD, by extracting a greatest common left (right) divisor. Furthermore, two coprime left (right) MFD of the same system differ only by a unimodular left (right) factor. Finally, if $\{N_L(s), D_L(s)\}$ and $\{N_R(s), D_R(s)\}$ are coprime, then $\deg\det D_L(s)$ and $\deg\det D_R(s)$ are minimum, and

$$\deg\det D_L(s) \;=\; \deg\det D_R(s) \;=\; n \;=\; \text{order of the system}$$

Lemma 1.1 : Invariance of row and column degrees

Let $\{N_L(s), D_L(s)\}$ be a coprime left MFD with $D_L(s)$ row reduced and let $\{N_R(s), D_R(s)\}$ be a coprime right MFD with $D_R(s)$ column reduced, then the row degrees of $D_L(s)$ are invariant and equal, except for possible permutations, to the *observability indices* $\{\nu_i\}$ of the system. Similarly, the column degrees of $D_R(s)$ are invariant and equal to the *controllability indices* $\{\mu_i\}$ of the system. Also,

$$\sum_{i=1}^{p} \partial r_i D_L = \sum_{i=1}^{p} \nu_i = n = \sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} \partial c_i D_R$$

See, *e.g*, Kailath [51], Chapter 6, for a proof of this result. Note that there is an infinite number of different coprime left (right) MFD's with $D_L(s)$ row reduced ($D_R(s)$ column reduced), which differ only by an unimodular left (right) factor.

Definition 1.12 : Pseudo-canonical left and right MFD

It can be shown, see Correa & Glover [52, 53] and Gevers & Wertz [54], that for any strictly proper system, and for any valid set of pseudo-observability indices $\{\rho_i\}$ and pseudo-controllability indices $\{\kappa_i\}$, there exist a unique left coprime MFD $\{N_L(s), D_L(s)\}$ called *pseudo-canonical left MFD* and a unique right coprime MFD $\{N_R(s), D_R(s)\}$ called *pseudo-canonical right MFD* such that:

$$\begin{cases} \partial c_i D_L &=& \rho_i \;\; \forall\, i \\ \Gamma_c[D_L] &=& I \\ \partial r_i D_L &\leq& \max(\rho_i, \rho_{\max} - 1) \;\; \forall\, i \\ \partial r_i N_L &<& \max(\rho_i, \rho_{\max} - 1) \;\; \forall\, i \end{cases} \qquad \begin{cases} \partial r_i D_R &=& \kappa_i \;\; \forall\, i \\ \Gamma_r[D_R] &=& I \\ \partial c_i D_R &\leq& \max(\kappa_i, \kappa_{\max} - 1) \;\; \forall\, i \\ \partial c_i N_R &<& \max(\kappa_i, \kappa_{\max} - 1) \;\; \forall\, i \end{cases} \tag{1.8}$$

Furthermore, it can be shown (see, *e.g.*, Correa & Glover [52, 53]) that the pseudo-observability canonical form $\{A_{poy}, B_{poy}, C_{poy}\}$ is the observability realization (as defined in Kailath [51]) of the pseudo-canonical left MFD and the pseudo-controllability canonical form $\{A_{pcy}, B_{pcy}, C_{pcy}\}$ is the controllability realization of the pseudo-canonical right MFD. Consequently,

$$
\begin{aligned}
D_{L_{ij}}(s) &= -\sum_{k=0}^{\rho_j-1} \alpha_{ijk} s^k \quad \text{if} \quad i \neq j \\
D_{L_{ii}}(s) &= s^{\rho_i} - \sum_{k=0}^{\rho_i-1} \alpha_{iik} s^k \\
D_{R_{ij}}(s) &= -\sum_{k=0}^{\kappa_i-1} \beta_{jik} s^k \quad \text{if} \quad i \neq j \\
D_{R_{ii}}(s) &= s^{\kappa_i} - \sum_{k=0}^{\kappa_i-1} \beta_{iik} s^k
\end{aligned}
$$

where the coefficients α_{ijk} and β_{ijk} are the same than in Definition 1.8 and 1.9. Also,

$$N_L(s) = \sum_{i=0}^{\rho_{max}-1} N_i s^i$$

$$\text{with} \quad \sum_{i=0}^{\rho_{max}-1} A^i_{poy} \; \text{block diag} \left\{ \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, (\rho_i \times 1), i = 1, \ldots, p \right\} N_i = B_{poy}$$

$$N_R(s) = \sum_{i=0}^{\kappa_{max}-1} N_i s^i$$

$$\text{with} \quad \sum_{i=0}^{\kappa_{max}-1} N^i \; \text{block diag} \left\{ \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}, (1 \times \kappa_i), i = 1, \ldots, m \right\} A^i_{pcy} = C_{pcy}$$

Definition 1.13 : Canonical left and right MFD

It was shown by Popov [58], Forney [59], Guidorzi [60], Beghelli & Guidorzi [61] and Kailath [51], that for any strictly proper system with observability indices $\{\nu_i\}$ and controllability indices $\{\mu_i\}$ there exists a unique coprime left MFD $\{N_L(s), D_L(s)\}$ and a unique coprime right MFD $\{N_R(s), D_R(s)\}$ called *canonical left echelon form* and *canonical right echelon form*, such that:

$$
\left\{
\begin{array}{lcl}
\partial c_i D_L &=& \nu_i \;\; \forall\, i \\
\Gamma_c[D_L] &=& I \\
\partial r_i D_L &=& \nu_i \;\; \forall\, i \\
\Gamma_r[D_L] &=& \begin{array}{l}\text{lower triangular} \\ \text{with unit diagonal}\end{array} \\
\partial r_i N_L &<& \nu_i \;\; \forall\, i
\end{array}
\right.
\qquad
\left\{
\begin{array}{lcl}
\partial r_i D_R &=& \mu_i \;\; \forall\, i \\
\Gamma_r[D_R] &=& I \\
\partial c_i D_L &=& \mu_i \;\; \forall\, i \\
\Gamma_c[D_R] &=& \begin{array}{l}\text{upper triangular} \\ \text{with unit diagonal}\end{array} \\
\partial c_i N_R &<& \mu_i \;\; \forall\, i
\end{array}
\right.
\qquad (1.9)
$$

The observability canonical form $\{A_{oy}, B_{oy}, C_{oy}\}$ is the observability realization (as defined in Kailath [51]) of the canonical left MFD and the controllability canonical form $\{A_{cy}, B_{cy}, C_{cy}\}$ is the controllability realization of the canonical right MFD. Consequently,

$$\begin{aligned}
D_{L_{ij}}(s) &= -\sum_{k=0}^{\nu_j-1} \alpha_{ijk} s^k \quad \text{if} \quad i \neq j \\
D_{L_{ii}}(s) &= s^{\nu_i} - \sum_{k=0}^{\nu_i-1} \alpha_{iik} s^k \\
D_{R_{ij}}(s) &= -\sum_{k=0}^{\mu_i-1} \beta_{jik} s^k \quad \text{if} \quad i \neq j \\
D_{R_{ii}}(s) &= s^{\mu_i} - \sum_{k=0}^{\mu_i-1} \beta_{iik} s^k
\end{aligned}$$

where the coefficients α_{ijk} and β_{ijk} are the same than in Definition 1.8 and 1.9 (some coefficients α_{ijk} and β_{ijk} are equal to zero). Also,

$$N_L(s) = \sum_{i=0}^{\nu_{max}-1} N_i s^i$$

$$\text{with} \quad \sum_{i=0}^{\nu_{max}-1} A_{oy}^i \ \text{block diag}\left\{ \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, (\nu_i \times 1), i = 1, \ldots, p \right\} N_i = B_{oy}$$

$$N_R(s) = \sum_{i=0}^{\mu_{max}-1} N_i s^i$$

$$\text{with} \quad \sum_{i=0}^{\mu_{max}-1} N^i \ \text{block diag}\left\{ \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}, (1 \times \kappa_i), i = 1, \ldots, m \right\} A_{cy}^i = C_{cy}$$

The pseudo-canonical MFD with indices identical to the indices of the canonical MFD will, of course, be identical to the canonical MFD.

It should be pointed out that the canonical $D_L(s)$ is always row-reduced, but the pseudo-canonical $D_L(s)$ is not always row-reduced. Similarly, the canonical $D_R(s)$ is always column-reduced, but the pseudo-canonical $D_R(s)$ is not always column-reduced. This is of some importance when identifying a pseudo-canonical left or right MFD since the constraints (1.8) do not guarantee that the estimated transfer function matrix is strictly proper (see Property 1.1).

Finally, it is interesting to note that the uniqueness of $D_L(s)$, and consequently of $N_L(s)$, only depends on the fact that its leading column coefficient matrix is the identity. Indeed,

suppose there existed another canonical or pseudo-canonical left MFD, $\{N_L^*(s), D_L^*(s)\}$, then there would exist a unimodular matrix $U(s)$ such that

$$\begin{array}{ll} & D_L^*(s) = U(s)D_L(s) \\ \Rightarrow & \lim_{s\to\infty} U(s) = \lim_{s\to\infty}(D_L^*(s)S_c^{-1}(s))(D_L(s)S_c^{-1}(s))^{-1} = I \\ \Rightarrow & U(s) = I \end{array}$$

This shows how the uniqueness of a left MFD is related to the knowledge of the observability indices or of a set of pseudo-observability indices and to the constraint that the *column* degrees (rather than the row degrees) be equal to them. Similarly, the uniqueness of $D_R(s)$, and consequently of $N_R(s)$, only depends on the fact that its leading row coefficient matrix is the identity.

In the case of the system given as example in Definition 1.6, 1.8, and 1.9

$$P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

the observability indices are $\{2,2\}$, the pseudo-observability indices $\{2,2\}$, $\{3,1\}$, and $\{1,3\}$, the controllability indices $\{3,1\}$, and the pseudo-controllability indices $\{3,1\}$. The canonical left MFD (which is also the pseudo-canonical left MFD with indices $\{2,2\}$) is given by

$$D_L(s) = \begin{bmatrix} s^2+2s+1 & s \\ s+1 & s^2+2s+2 \end{bmatrix} \quad \Gamma_c[D_L] = I \quad \Gamma_r[D_L] = I \qquad N_L(s) = \begin{bmatrix} s & -2s-2 \\ 1 & s-1 \end{bmatrix}$$

The pseudo-canonical left MFD with indices $\{3,1\}$ is

$$\begin{aligned} D_L(s) &= \begin{bmatrix} s^3+4s^2+4s+1 & -2 \\ s^2+2s+1 & s \end{bmatrix} \quad \Gamma_c[D_L] = I \quad \Gamma_r[D_L] = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \\ N_L(s) &= \begin{bmatrix} s^2+2s-1 & -2s^2-7s-3 \\ s & -2s-2 \end{bmatrix} \end{aligned}$$

The pseudo-canonical left MFD with indices $\{1,3\}$ is

$$\begin{aligned} D_L(s) &= \begin{bmatrix} s+1 & s^2+2s+2 \\ 0 & s^3+3s^2+3s+2 \end{bmatrix} \quad \Gamma_c[D_L] = I \quad \Gamma_r[D_L] = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \\ N_L(s) &= \begin{bmatrix} 1 & s-1 \\ 1 & s^2+2s+1 \end{bmatrix} \end{aligned}$$

The canonical right MFD (which is also the pseudo-canonical right MFD with indices $\{3,1\}$) is given by

$$D_R(s) = \begin{bmatrix} s^3+3s^2+3s+2 & 1 \\ 0 & s+1 \end{bmatrix} \quad \Gamma_r[D_R] = I \quad \Gamma_c[D_R] = I \qquad N_R(s) = \begin{bmatrix} s^2+s & -2 \\ 1 & 1 \end{bmatrix}$$

Chapter 2

Multivariable adaptive control

2.0 Introduction

In this work, we are mostly concerned with *parametric* estimation and *parametric* adaptive control of deterministic continuous-time linear systems. There are two different types of parametric adaptive control schemes: direct and indirect. Indirect schemes estimate the parameters of the transfer function of the process. Then, the parameters of the controller are updated based on these estimates. The update mechanism of the controller is often a complex nonlinear transformation between the estimated parameters and the controller parameters which depends on the technique used to design the controller. In direct adaptive control, the parameters of the controller are estimated directly. This is often possible when the system transfer function can be reparameterized in terms of the regulator parameters. In that case, the estimator updates directly the regulator parameters and the adjustment mechanism between the estimated parameters and the controller parameters is simply the identity transformation.

The main direct adaptive control schemes for continuous-time linear systems are the direct model reference adaptive control (MRAC) scheme and the direct adaptive pole placement (PPAC) scheme. In the MRAC scheme, the desired output of the closed-loop system is specified through a reference model, and the adaptive controller tries to make the plant output match the reference model output. There exist two different ways to implement a direct MRAC scheme. The first one uses for identification the output error (the difference between the output of the plant and the output of the model) and assumes that the reference model is strictly positive real (SPR), (*cf.* Definition 1.5). The other one uses the input error (the difference between the the inverse of the model applied to the output of the plant and the reference input) and does not require such assumption. Another advantage of input error schemes is the fact that they can be used in conjunction with recursive least-squares parameter estimation algorithms, while output error schemes are limited to gradient type algorithms whose convergence is very slow when there is a large number of parameters to estimate. The main drawback to the MRAC scheme is the fact that it is applicable only to minimum phase systems (systems whose zeros are

in the open left-half plane). In the PPAC scheme, only the poles of the closed-loop system are assigned adaptively to predefined locations. The PPAC scheme can be applied to minimum and non-minimum phase systems. As opposed to direct adaptive control, indirect adaptive control allows for the use of almost any type of control method and any type of recursive identifier. On the other hand, stability may be difficult, if not impossible, to guarantee. In this chapter, we present continuous-time direct MRAC and PPAC algorithms and also a recursive identifier that could be used in an indirect adaptive control scheme. The algorithms presented in this chapter are similar to algorithms that are found in, *e.g.*, Elliot & Wolovich [6, 11], Elliot *et al.* [32], and Das [7]. After some preliminaries, we first show how to define a model reference controller and a pole placement controller assuming that the plant is known. Then, we show how to introduce parameter estimation and adaptation if the plant is unknown. A recursive identifier that could be used in an indirect scheme is also presented. Then, we give the available stability results for the two direct schemes (MRAC and PPAC). Finally, a comparison is made between schemes based on the required *a priori* information.

2.1 Preliminaries

As mentionned in Section 1.1, $P(s)$ denotes the transfer function of a linear time invariant operator, P. Then, $P[u]$ denotes the output of the operator in the time domain with input $u(t)$. Furthermore, the dependency in s of a transfer function or the dependency in t of a signal will be often omitted when there is no ambiguity.

A polynomial in s is called *monic* if the coefficient of the highest power in s is 1 and *Hurwitz* if its roots lie in the open left-half plane. Rational transfer functions are called *stable* if their denominator polynomial is Hurwitz and *minimum phase* if their numerator polynomial is Hurwitz. A matrix transfer function is stable if all its elements are stable and is minimum phase if its inverse is stable.

We will denote by $\nu_{\max}$ the maximum of the observability indices $\{\nu_i\}$ (*cf.* Definition 1.6). Similarly, $\mu_{\max}$ will denote the maximum of the controllability indices $\{\mu_i\}$ (*cf.* Definition 1.7).

The following lemmas will be useful to build our direct multivariable adaptive controllers.

Lemma 2.1 : Bezout identity

Two polynomial matrices $N(s)$ and $D(s)$ are left coprime iff:

$\exists$ two polynomial matrices $U(s)$ and $V(s)$ such that:

$$N(s)U(s) + D(s)V(s) = I$$

and right coprime iff:

$\exists$ two polynomial matrices $U(s)$ and $V(s)$ such that:

$$U(s)N(s) + V(s)D(s) = I$$

A proof can be found in Kailath [51], Chapter 6, pp. 379.

Lemma 2.2 : Polynomial matrix division
Let $N_R(s), N_L(s) \in \Re^{p\times m}[s]$, $D_R(s) \in \Re^{m\times m}[s]$, $D_L(s) \in \Re^{p\times p}[s]$ with $D_R(s)$ and $D_L(s)$ non-singular. There exist unique matrices $Q_R(s)$, $R_R(s)$, $Q_L(s)$, $R_L(s)$,$\in \Re^{p\times m}[s]$ such that:

$$N_R(s) = Q_R(s)D_R(s) + R_R(s) \quad \text{and} \quad R_R(s)D_R(s)^{-1} \text{ is strictly proper}$$

$$N_L(s) = D_L(s)Q_L(s) + R_L(s) \quad \text{and} \quad D_L(s)^{-1}R_L(s) \text{ is strictly proper}$$

We will call $Q_R(s)$ and $Q_L(s)$ the quotient, and $R_R(s)$ and $R_L(s)$ the remainder, of the matrix divisions $N_R(s)D_R^{-1}(s)$ and $D_L^{-1}(s)N_L(s)$. It can be verified using Property 1.1 that

$$\partial c_i R_R < \partial c_i D_R \quad \text{and} \quad \partial r_i R_L < \partial r_i D_L$$

Furthermore, if $D_R(s)$ is column reduced and $D_L(s)$ row reduced, there exist unique matrices $Q_R(s)$, $R_R(s)$, $Q_L(s)$, $R_L(s)$,$\in \Re^{p\times m}[s]$ such that:

$$N_R(s) = Q_R(s)D_R(s) + R_R(s) \quad \text{and} \quad \partial c_i R_R < \partial c_i D_R$$

$$N_L(s) = D_L(s)Q_L(s) + R_L(s) \quad \text{and} \quad \partial r_i R_L < \partial r_i D_L$$

The proof of this lemma is in Kailath [51], Chapter 6, pp. 388-390.

2.2 Control algorithm

2.2.1 Controller structure

The following controller structure is considered (see Fig. 2.1)

$$\begin{aligned} r &= M_0[r_0] \\ u &= C_0 r + \Lambda^{-1}C[u] + \Lambda^{-1}D[y_p] \end{aligned} \tag{2.1}$$

where u is the vector of inputs of the system, y_p the outputs, r the reference signals, $C_0 \in \Re^{m\times m}$ is nonsingular, $\Lambda(s), C(s) \in \Re^{m\times m}[s]$, $D(s) \in \Re^{m\times p}[s]$, and $M_0(s) \in \Re^{m\times m}(s)$. $M_0(s)$ is a proper stable transfer function matrix and $\Lambda(s)$ is a Hurwitz diagonal matrix, $\Lambda(s) = \mathrm{diag}\{\lambda_i(s)\}$, such

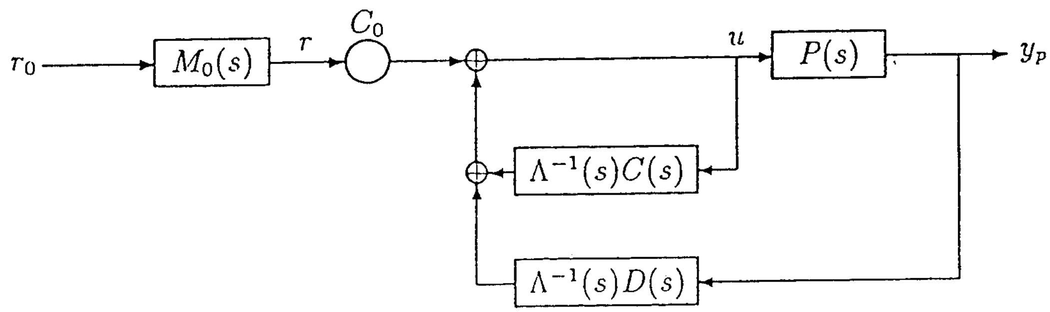

Figure 2.1: Controller structure

that $\Lambda^{-1}D$ is proper and $\Lambda^{-1}C$ is strictly proper. It is sometimes useful to rewrite the control law (2.1) as

$$u = r + \Lambda^{-1}\bar{C}[u] + \Lambda^{-1}\bar{D}[y_p] \tag{2.2}$$

where

$$\begin{aligned} \Lambda^{-1}\bar{D} &= C_0^{-1}\Lambda^{-1}D \quad \text{proper} \\ \Lambda^{-1}\bar{C} &= C_0^{-1}\Lambda^{-1}C + (I - C_0^{-1}) \quad \text{proper} \\ \lim_{s\to\infty} \Lambda^{-1}\bar{C} &= \bar{C}_{0_\infty} = (I - C_0^{-1}) \end{aligned}$$

However, since $\Lambda^{-1}\bar{C}$ is only proper, one must still use for implementation the following equivalent control law

$$u = (I - \bar{C}_{0_\infty})^{-1}(r + \Lambda^{-1}(\bar{C} - \Lambda\bar{C}_{0_\infty})[u] + \Lambda^{-1}\bar{D}[y_p]) \tag{2.3}$$

By combining the equation of the plant, $y_p = P[u]$, with the equation of the controller (2.1) or (2.2), the output y_p can be expressed as

$$\begin{aligned} y_p &= N_R((\Lambda - C)D_R - DN_R)^{-1}\Lambda C_0[r] \\ &= N_R((\Lambda - \bar{C})D_R - \bar{D}N_R)^{-1}\Lambda[r] \end{aligned} \tag{2.4}$$

where $\{N_R, D_R\}$ is a right MFD of $P(s)$.

2.2.2 Model reference

Assumptions

(A1) Plant assumptions
The plant is described by a square, nonsingular, strictly proper, and minimum phase transfer function matrix $P(s) \in \Re^{p\times p}(s)$.

(A2) Model assumptions
The reference model $M(s) = H(s)M_0(s)$, where $H(s)$ is the Hermite normal form of $P(s)$ (see Definition 1.4), and $M_0(s) \in \Re^{p\times p}(s)$ is a proper stable transfer function matrix.

(A3) Reference input assumptions
The reference input, $r_0(t)$, is piecewise continuous and belongs to L_∞ $\forall\, t > 0$.

The model output, y_m, is defined as:

$$y_m = HM_0[r_0] = H[r] \tag{2.5}$$

Now, we show that there exist controller parameters values such that the output of the plant y_p matches the output of the reference model y_m.

Proposition 2.1 : Model reference matching equality
If $\partial\lambda_i = \nu - 1$ and $\nu \geq \nu_{\max}$, then there exist $C_0^* \in \Re^{p\times p}$, $C^*(s)$, $D^*(s) \in \Re^{p\times p}[s]$, solution of the Diophantine equation:

$$N_R[(\Lambda - C^*)D_R - D^*N_R]^{-1}\Lambda C_0^* = H \tag{2.6}$$

such that model matching is achieved (*i.e.*, the transfer function from r to y_p is $H(s)$), $\Lambda^{-1}D^*$ is proper, and $\Lambda^{-1}C^*$ is strictly proper. In particular, $C_0^* = K_p^{-1}$ nonsingular, $\partial D^* \leq \nu_{\max} - 1$, and $\partial r_i C^* < \partial\lambda_i$. Similarly, $\exists \bar{C}^*(s)$ and $\bar{D}^*(s) \in \Re^{p\times p}[s]$ such that model matching is achieved, $\Lambda^{-1}\bar{D}^*$ and $\Lambda^{-1}\bar{C}^*$ are proper, and $\bar{C}^*_{0_\infty} = I - K_p$.

Proof
The matching equality (2.6) is equivalent to

$$C_0^* H^{-1} P = I - \Lambda^{-1}C^* - \Lambda^{-1}D^* P \tag{2.7}$$

Now, let $\{N_L, D_L\}$ be a left coprime MFD of $P(s)$ with D_L row reduced ($\partial D_L(s) \leq \nu_{\max}$, *cf.* Lemma 1.1). Using the polynomial matrix division (Lemma 2.2), divide $\Lambda K_p^{-1} H^{-1}$ on the right by D_L, then $\exists$ $Q(s)$, $R(s) \in \Re^{p\times p}[s]$ such that

$$\Lambda K_p^{-1} H^{-1} = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

and $\partial c_i R < \partial c_i D_L \leq \nu_{\max}$. Then, let

$$D^* = -R = QD_L - \Lambda K_p^{-1} H^{-1} \qquad C^* = \Lambda - QN_L \qquad C_0^* = K_p^{-1}$$

It can be easily verified that the given C_0^*, C^*, D^* solve the matching equality (2.7). Furthermore, since $\partial\lambda_i = \nu - 1$, $\Lambda^{-1}D^*$ is proper. On the other hand,

$$\lim_{s\to\infty} \Lambda^{-1}C^* = \lim_{s\to\infty}(I - K_p^{-1}H^{-1}P - \Lambda^{-1}D^*P) = I - I = 0$$

so that $\Lambda^{-1}C^*$ is strictly proper and $\partial r_i C^* \leq \partial\lambda_i - 1$, *e.g.*, $\partial r_i C^* \leq \nu_{\max} - 2$ if $\partial\lambda_i = \nu_{\max} - 1$.

Similarly, dividing ΛH^{-1} on the right by D_L, then $\exists\ Q(s),\ R(s) \in \Re^{p\times p}[s]$ such that

$$\Lambda H^{-1} = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

Let

$$\bar{D}^* = \bar{Q}D_L - \Lambda H^{-1} \qquad \bar{C}^* = \Lambda - \bar{Q}N_L$$

then $\partial\bar{D}^* \leq \nu_{\max} - 1$, $\partial r_i\bar{C}^* \leq \partial\lambda_i$, $\bar{C}^*_{0_\infty} = I - K_p$, and the given $\bar{C}^*$ and $\bar{D}^*$ are a solution of the matching equality

$$N_R[(\Lambda - \bar{C}^*)D_R - \bar{D}^*N_R]^{-1}\Lambda = H$$

□

Note that the controller parameters C_0^*, C^*, and D^* (or $\bar{C}^*$ and $\bar{D}^*$) are not uniquely defined. We will show in the next chapter (Chapter 3) how to guarantee uniqueness of the controller parameters.

2.2.3 Pole placement

Assumptions

(A1) Plant assumptions
The plant is described by a strictly proper, transfer function matrix $P(s) \in \Re^{p\times m}(s)$.

(A2) Model assumptions
The pole placement objective is equivalent to assuming a reference model $M(s) = N_R(s)D_M^{-1}(s)$ where $\{N_R(s), D_R(s)\}$ is the canonical right MFD of $P(s)$ (*cf.* Definition 1.13), and $D_M(s) = \mathrm{diag}\{d_i(s)\} \in \Re^{m\times m}[s]$, with $d_i(s)$ monic, Hurwitz, and $\partial d_i(s) = \mu_i$.

(A3) Reference input assumptions
The reference input, $r_0(t)$, is piecewise continuous and bounded on $\forall\, t > 0$.

The model output is defined as

$$y_m = N_R D_M^{-1}[r] \tag{2.8}$$

Similarly, to the model reference case, we show that there exist controller parameters values such that y_p matches y_m.

Proposition 2.2 : Pole placement matching equality

If $\partial\lambda_i = \nu - 1$ and $\nu \geq \nu_{\max}$, $\exists C_0^* \in \Re^{m\times m}$, $C^*(s) \in \Re^{m\times m}[s]$, $D^*(s) \in \Re^{m\times p}[s]$, solution of the Diophantine equation:

$$[(\Lambda - C^*)D_R - D^* N_R]^{-1}\Lambda C_0^* = D_M^{-1} \tag{2.9}$$

such that model matching is achieved, $\Lambda^{-1}D^*$ is proper, and $\Lambda^{-1}C^*$ is strictly proper. In particular, $C_0^* = \Gamma_c[D_R]$ nonsingular, $\partial D^* \leq \nu_{\max} - 1$, and $\partial r_i C^* < \partial\lambda_i$. Similarly, $\exists\ \bar{C}^* \in \Re^{m\times m}[s]$ and $\bar{D}^* \in \Re^{m\times p}[s]$ such that model matching is achieved, $\Lambda^{-1}\bar{D}^*$ and $\Lambda^{-1}\bar{C}^*$ are proper, and $\bar{C}_{0_\infty}^* = I - \Gamma_c[D_R]^{-1}$.

Proof

The proof is similar to the model reference case. The matching equality (2.9) is equivalent to

$$C_0^* D_M D_R^{-1} = I - \Lambda^{-1}C^* - \Lambda^{-1}D^* P \tag{2.10}$$

Since $\{N_R, D_R\}$ are coprime, by the Bezout identity (Lemma 2.1), $\exists\ U_L^* \in \Re^{m\times p}[s]$, $V_L^* \in \Re^{m\times m}[s]$ such that

$$U_L^* N_R + V_L^* D_R = I \quad \Leftrightarrow \quad U_L^* P + V_L^* = D_R^{-1}$$

Now, let $\{N_L, D_L\}$ be a left coprime MFD of $P(s)$ with D_L row reduced. Using the polynomial matrix division (Lemma 2.2), divide $\Lambda\Gamma_c[D_R]D_M U_L^*$ on the right by D_L, then $\exists\ Q(s), R(s) \in \Re^{m\times p}[s]$ such that

$$\Lambda\Gamma_c[D_R]D_M U_L^* = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

and $\partial c_i R < \partial c_i D_L \leq \nu_{\max}$. Then, let

$$D^* = -R = QD_L - \Lambda\Gamma_c[D_R]D_M U_L^* \qquad C^* = \Lambda - QN_L - \Lambda\Gamma_c[D_R]D_M V_L^* \qquad C_0^* = \Gamma_c[D_R]$$

It is easy to verify that the given C_0^*, C^*, D^* solve the matching equality (2.10). Furthermore, since $\partial\lambda_i \geq \nu - 1$, $\Lambda^{-1}D^*$ is proper. On the other hand

$$\lim_{s\to\infty} \Lambda^{-1}C^* = \lim_{s\to\infty}(I - \Gamma_c[D_R]D_M(U_L^* P + V_L^*) - \Lambda^{-1}D^* P) = I - I = 0$$

so that $\Lambda^{-1}C^*$ is strictly proper and $\partial r_i C^* \leq \partial\lambda_i - 1$.

Similarly, dividing $\Lambda D_M U_L^*$ on the right by D_L, then $\exists\ Q(s), R(s) \in \Re^{m\times p}[s]$ such that

$$\Lambda D_M U_L^* = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

Let

$$\bar{D}^* = \bar{Q}D_L - \Lambda D_M U_L^* \qquad \bar{C}^* = \Lambda - \bar{Q}N_L - \Lambda D_M V_L^*$$

then $\partial\bar{D}^* \leq \nu_{\max}-1$, $\partial r_i\bar{C}^* \leq \partial\lambda_i$, $\bar{C}^*_{0_\infty} = I - \Gamma_c[D_R]^{-1}$, and the given $\bar{C}^*$ and $\bar{D}^*$ are a solution of the matching equality

$$[(\Lambda - \bar{C}^*)D_R - \bar{D}^* N_R]^{-1}\Lambda = D_M^{-1}$$

□

Note that the controller parameters C_0^*, C^*, and D^* (or $\bar{C}^*$ and $\bar{D}^*$) are not uniquely defined. We will show in the next chapter (Chapter 3) how to guarantee uniqueness of the controller parameters.

2.3 Adaptation

In order to implement adaptation in the model reference and pole placement designs, one must estimate C_0^*, C^*, D^*, (or $\bar{C}^*, \bar{D}^*$), which satisfy the matching equality. This section presents a direct MRAC algorithm, a direct adaptive pole placement algorithm and also a recursive identifier that could be used in an indirect scheme.

2.3.1 Model reference adaptive control

The matching equality (2.6) is equivalent to

$$\begin{aligned} I &= C_0^* H^{-1}P + \Lambda^{-1}C^* + \Lambda^{-1}D^*P \\ &= H^{-1}P + \Lambda^{-1}\bar{C}^* + \Lambda^{-1}\bar{D}^*P \end{aligned} \tag{2.11}$$

Define $L(s) = \mathrm{diag}\{l(s)\}$, with $l(s)$ Hurwitz, $\partial l(s) \geq d$, where d is the maximum degree of all elements of $H^{-1}(s) = \xi(s)$. Then, multiplying both sides of (2.11) by L^{-1} and applying both transfer function matrices to u leads to

$$L^{-1}[u] = C_0^*(HL)^{-1}[y_p] + L^{-1}(\Lambda^{-1}C^*[u] + \Lambda^{-1}D^*[y_p]) \tag{2.12}$$

If $H(s)$ is known *a priori*, then (2.12) is an equation where the unknown parameters appear linearly. Indeed, given Proposition 2.1, there exist matrices $C_1^*, \ldots, C_{\nu-1}^*$, $D_1^*, \ldots, D_\nu^* \in \Re^{p\times p}$ such that

$$\begin{aligned} \Lambda^{-1}C^* &= \sum_{i=1}^{\nu-1} C_i^* \frac{s^{(i-1)}}{\lambda(s)} \\ \Lambda^{-1}D^* &= D_\nu^* + \sum_{i=1}^{\nu-1} D_i^* \frac{s^{(i-1)}}{\lambda(s)} \end{aligned}$$

The matrix of unknown controller parameters is

$$\theta^{*T} = \left[\begin{array}{cccccc} C_0^* & \dots & C_{\nu-1}^* & D_1^* & \dots & D_\nu^* \end{array} \right]$$

and the regressor vector

$$\begin{aligned} \psi^T = \Big[& (HL)^{-1}[y_p]^T \quad (\Lambda L)^{-1}[u]^T \quad \dots \quad s^{(\nu-2)}(\Lambda L)^{-1}[u]^T \\ & (\Lambda L)^{-1}[y_p]^T \quad \dots \quad s^{(\nu-2)}(\Lambda L)^{-1}[y_p]^T \quad L^{-1}[y_p]^T \Big] \end{aligned} \tag{2.13}$$

so that (2.12) can be rewritten as

$$L^{-1}[u] = \theta^{*^T}\psi \tag{2.14}$$

Then the following error equation can be derived

$$e_2 = \theta^T\psi - L^{-1}[u] = (\theta^T - \theta^{*^T})\psi = \phi^T\psi \tag{2.15}$$

where θ is the estimate of θ^* and ϕ is the parameter error. At this point, θ may be estimated using a standard linear estimation algorithm. We will consider the normalized least-squares algorithm with covariance resetting (see section 2.4).

The error equation is equivalent to the one presented in Elliot & Wolovich [6, 11] and included in Sastry & Bodson [3]. It requires the *a priori* knowledge of the Hermite normal form $H(s)$ and an upper bound ν on the observability index $\nu_{\max}$. Then, the number of parameters to identify, N_θ, is $2p^2\nu$.

2.3.2 Extended model reference adaptive control

If the normal Hermite form is not known, but the coefficients $\{r_i\}$ are, it is still possible to identify the controller parameters. Using the parameters $\bar{C}^*$, $\bar{D}^*$, the matching equality (2.6) is equivalent to

$$I = H^{-1}P + \Lambda^{-1}\bar{C}^* + \Lambda^{-1}\bar{D}^*P \tag{2.16}$$

Then, using the properties of $H^{-1}(s)$ (see Definition 1.4) and applying the transfer function matrices to u leads to

$$L^{-1}[u] - L^{-1}\Delta[y_p] = L^{-1}(\Sigma^*F\Delta[y_p] + \Lambda^{-1}\bar{C}^*[u] + \Lambda^{-1}\bar{D}^*[y_p]) \tag{2.17}$$

where L^{-1} is as previously defined. By defining θ^* and ψ accordingly, this is equivalent to

$$L^{-1}[u] - L^{-1}\Delta[y_p] = \theta^{*^T}\psi \tag{2.18}$$

so that standard linear estimation techniques can again be used.

This algorithm is essentially the one presented in Das [7] and Dion *et al.* [50]. In this algorithm, the number of parameters $N_\theta \geq 2p^2\nu$.

2.3.3 Adaptive pole placement

The matching equality (2.9) is equivalent to

$$\begin{aligned} I &= C_0^* D_M D_R^{-1} + \Lambda^{-1} C^* + \Lambda^{-1} D^* P \\ &= D_M D_R^{-1} + \Lambda^{-1} \bar{C}^* + \Lambda^{-1} \bar{D}^* P \end{aligned} \tag{2.19}$$

Unfortunately, D_R being unknown, the previous scheme cannot be directly used. However, since $\{N_R, D_R\}$ is coprime, by the Bezout identity (Lemma 2.1), $\exists\ U_L^* \in \Re^{m\times p}[s]$, $V_L^* \in \Re^{m\times m}[s]$ such that

$$V_L^* D_R + U_L^* N_R = I \quad \Leftrightarrow \quad V_L^* + U_L^* P = D_R^{-1} \tag{2.20}$$

Note that U_L^* and V_L^* are generally not unique. It can also be shown that there exist U_L^* and V_L^* such that $\partial U_L^* \leq \nu_{\max} - 1$ and $\partial V_L^* < \partial U_L^*$.

Proposition 2.3 : Degree of the elements of the Bezout identity

Given a right coprime MFD $\{N_R, D_R\}$, $\exists$ matrices $U_L^* \in \Re^{m\times p}[s]$, $V_L^* \in \Re^{m\times m}[s]$ such that

$$V_L^* D_R + U_L^* N_R = I \quad \text{and} \quad \partial U_L^* \leq \nu_{\max} - 1$$

Furthermore, $\partial V_L^* < \partial U_L^*$.

Proof

Since $\{N_R, D_R\}$ are coprime, we know by the Bezout identity (Lemma 2.1) that there exists a solution $\bar{U}_L \in \Re^{m\times p}[s]$, $\bar{V}_L \in \Re^{m\times m}[s]$ such that

$$\bar{V}_L D_R + \bar{U}_L N_R = I$$

Divide $\bar{U}_L$ on the right by D_L, where $\{N_L, D_L\}$ is a left MFD with D_L row reduced ($\partial D_L \leq \nu_{\max}$, *cf.* Lemma 1.1), then by the polynomial matrix division (Lemma 2.2) $\exists\ Q_L$ and $U_L^* \in \Re^{m\times p}[s]$ such that

$$\bar{U}_L = Q_L D_L + U_L^* \quad \text{and} \quad U_L^* D_L^{-1} \text{ is strictly proper}$$

and $\partial U_L^* \leq \nu_{\max} - 1$. Now let

$$V_L^* = \bar{V}_L + Q_L N_L$$

It can be easily verified that U_L^*, V_L^* is also a solution of the Bezout identity. Furthermore, from (2.20), we have that

$$V_L^* = D_R^{-1} - U_L^* P$$

with P strictly proper, so that $\partial V_L^* < \partial U_L^*$ (*i.e.*, $\partial V_L^* \leq \nu_{\max} - 2$ if $\partial c_i U_L^* \leq \nu_{\max} - 1$). □

Now, define $L(s) = \text{diag}\{l(s)\}$, with $l(s)$ Hurwitz, $\partial l(s) = \nu - 1$, and $\nu \geq \nu_{\max}$. Then using (2.20), multiplying both sides of (2.19) by $(D_M L)^{-1}$, and applying both transfer function matrices to u, leads to

$$(D_M L)^{-1}[u] = L^{-1}((\Lambda D_M)^{-1} \bar{C}^*[u] + (\Lambda D_M)^{-1} \bar{D}^*[y_p] + V_L^*[u] + U_L^*[y_p]) \tag{2.21}$$

This is an equation where the unknown parameters appear linearly. However, to achieve the pole placement objective, we must guarantee that, for all inputs u, equation (2.21) has a solution which is also the solution of the pole placement matching equality (2.19) and the Bezout identity (2.20).

Proposition 2.4 : Equivalence of (2.21) with (2.19) and (2.20)
If

$$\partial\lambda_i = \mu_{\max} - \mu_i + \nu - 1$$

then there exists a solution $\bar{C}^*, \bar{D}^*$ of the matching equality (2.19) such that

$$\begin{aligned} \partial c_i \bar{D}^* &\leq \nu_{\max} - 1 \leq \nu - 1 \\ \partial r_i \bar{C}^* &\leq \mu_{\max} - \mu_i + \nu - 1 \\ \partial c_i \bar{C}^* &\leq \mu_{\max} - \mu_i + \nu - 2 \\ \bar{C}^*_{\infty_{ij}} &= (I - \Gamma_c[D_R]^{-1})_{ij} = 0 \quad \text{if} \quad i \geq j \end{aligned} \tag{2.22}$$

Furthermore, under these constraints, any solution of (2.21) valid for all u is also a solution of (2.19) and (2.20).

Proof
From Proposition 2.2, we know that there exists a solution $\bar{C}^*, \bar{D}^*$ to the matching equality (2.19) such that

$$\begin{aligned} \partial c_i \bar{D}^* &\leq \nu_{\max} - 1 \leq \nu - 1 \\ \partial r_i \bar{C}^* &\leq \partial\lambda_i \\ \lim_{s\to\infty} \Lambda^{-1}\bar{C}^* &= \bar{C}^*_{0_\infty} = I - \Gamma_c[D_R] \quad \Rightarrow \quad \bar{C}^*_{0_{\infty_{ij}}} = 0 \quad \text{if} \quad i \geq j \end{aligned}$$

If $\partial\lambda_i = \mu_{\max} - \mu_i + \nu - 1$ then $\partial r_i \bar{C}^* \leq \mu_{\max} - \mu_i + \nu - 1$. Furthermore, $\partial r_i(\Lambda D_M) = \mu_{\max} + \nu - 1$ and $(\Lambda D_M)^{-1}\bar{D}^* N_R$ is strictly proper. Consequently, using (2.19)

$$\lim_{s\to\infty} (\Lambda D_M)^{-1}\bar{C}^* D_R = \lim_{s\to\infty} [D_M^{-1} D_R - I - (\Lambda D_M)^{-1}\bar{D}^* N_R] = I - I = 0$$

So that $(\Lambda D_M)^{-1}\bar{C}^* D_R$ is strictly proper and $\partial c_i \bar{C}^* \leq \mu_{\max} - \mu_i + \nu - 2$.
Finally, if equation (2.21) must be verified $\forall\, u$

$$(D_M L)^{-1} = L^{-1}((\Lambda D_M)^{-1}\bar{C}^* + (\Lambda D_M)^{-1}\bar{D}^* P + V_L^* + U_L^* P)$$

or equivalently

$$D_M^{-1}(D_R - D_M) - (\Lambda D_M)^{-1}\bar{C}^* D_R - (\Lambda D_M)^{-1}\bar{D}^* N_R = V_L^* N_R + U_L^* D_R - I$$

which is strictly proper on one side and polynomial on the other, so that both sides must be identically zero, or equivalently

$$\begin{aligned} I &= D_M D_R^{-1} + \Lambda^{-1}\bar{C}^* + \Lambda^{-1}\bar{D}^* P \\ I &= V_L^* N_R + U_L^* D_R \end{aligned}$$

Therefore, a solution of (2.21) for all inputs u is also a solution of (2.19) and (2.20). $\square$

Now, let us assume, for simplicity, that $\{\lambda_i(s)\}$ and $\{d_i(s)\}$ are such that $\lambda_i(s)d_i(s) = \gamma(s)\ \ \forall i$. Then $\Gamma(s) = \mathrm{diag}\{\gamma(s)\}$ is stable with $\partial\gamma(s) = \mu_{\max} + \nu - 1$. Based on Propositions 2.2, 2.3, and 2.4, we can define matrices $C_1^*, \ldots, C^*_{\mu_{\max}-\mu_{\min}+\nu-1}$, $D_1^*, \ldots, D_\nu^*$, $V_1^*, \ldots, V_{\nu-1}^*$, $U_1^*, \ldots, U_\nu^*$, $\in \Re^{p\times p}$ such that

$$\begin{aligned} (\Gamma L)^{-1}\bar{C}^* &= \sum_{i=1}^{\mu_{\max}-\mu_{\min}+\nu-1} C_i^* \frac{s^{(i-1)}}{l(s)\gamma(s)} \\ (\Gamma L)^{-1}\bar{D}^* &= \sum_{i=1}^{\nu} D_i^* \frac{s^{(i-1)}}{l(s)\gamma(s)} \\ L^{-1}V_L^* &= \sum_{i=1}^{\nu-1} V_i^* \frac{s^{(i-1)}}{l(s)} \\ L^{-1}U_L^* &= \sum_{i=1}^{\nu} U_i^* \frac{s^{(i-1)}}{l(s)} \end{aligned}$$

Note that, given the conditions (2.22) on $\bar{C}^*$, the matrices C_i^* with $i > (\nu - 1)$ contain several elements equal to zero. The matrix of unknown controller parameters is

$$\theta^{*^T} = \left[\begin{array}{cccccccccccc} C_1^* & \ldots & C^*_{\mu_{\max}-\mu_{\min}+\nu-1} & D_1^* & \ldots & D_\nu^* & V_1^* & \ldots & V_{\nu-1}^* & U_1^* & \ldots & U_\nu^* \end{array}\right]$$

and the regressor vector

$$\begin{aligned} \psi^T = \Big[& (\Gamma L)^{-1}[u]^T \ \ \ldots \ \ s^{(\mu_{\max}-\mu_{\min}+\nu-2)}(\Gamma L)^{-1}[u]^T \ \ (\Gamma L)^{-1}[y_p]^T \ \ \ldots \\ & s^{(\nu-1)}(\Gamma L)^{-1}[y_p]^T \ \ L^{-1}[u]^T \ \ \ldots \ \ s^{(\nu-1)}L^{-1}[u]^T \ \ L^{-1}[y_p]^T \ \ \ldots \ \ s^{(\nu-1)}L^{-1}[y_p]^T \Big] \end{aligned} \tag{2.23}$$

so that (2.21) can be rewritten as

$$L^{-1}D_M^{-1}(u) = \theta^{*^T}\psi \tag{2.24}$$

and an error equation can be derived

$$e_2 = \theta^T\psi - L^{-1}D_M^{-1}(u) = (\theta^T - \theta^{*^T})\psi = \phi^T\psi \tag{2.25}$$

where θ is the estimate of θ^* and may be estimated using a standard linear estimation algorithm. Note that, with the conditions (2.22), $(I - \bar{C}_{0_\infty})$ is upper triangular with one's on the diagonal and is guaranteed to remain nonsingular. Furthermore, since $(I - \bar{C}_{0_\infty})^{-1}$ is an upper triangular matrix, it can be easily inverted.

The pole placement algorithm requires the *a priori* knowledge of the controllability indices $\{\mu_i\}$ and an upper bound ν on the observability index $\nu_{\max}$. The number of parameters to identify, N_θ, is

$$2m^2(\nu - 1) + 2mp\nu \leq N_\theta \leq 2m^2(\nu - 1) + 2mp\nu + m(m\mu_{\max} - n)$$

This algorithm is similar to the one presented in Elliot *et al.* [32]. However, their algorithm does not match a fixed model $N_R D_M^{-1}$ but a member of the class of models $N_R U D_M^{-1}$ where $U(s)$ is an unimodular polynomial matrix.

2.3.4 Recursive identification

A recursive identifier is reviewed here. Such an identifier could be used in an indirect adaptive control scheme. Assume that $\{N_L, D_L\}$ is a left MFD of a strictly proper transfer matrix, $P(s) \in \Re^{p\times m}$ with D_L row reduced, and $\Lambda(s) = \mathrm{diag}\{\lambda_i(s)\}$, with $\partial\lambda_i = \nu \geq \nu_{\max}$. Then

$$\Lambda^{-1} N_L = \Lambda^{-1} D_L P \tag{2.26}$$

Applying (2.26) to u leads to

$$y_p = \Lambda^{-1} N_L[u] - \Lambda^{-1}(D_L - \Lambda)[y_p] \tag{2.27}$$

which is an equation where the unknown parameters appear linearly. By properly defining θ^* the matrix of unknown parameters and ψ the regressor vector, a linear error equation can be derived

$$e_2 = \theta^T\psi - y_p = (\theta^T - \theta^{*T})\psi = \phi^T\psi \tag{2.28}$$

The minimum necessary *a priori* information is $\nu \geq \nu_{\max}$. Indeed, if only ν is known, there exists a possibly noncoprime left MFD, $\{N_L, D_L\}$ such that $\partial N_L \leq \nu - 1$ and $\partial(D_L - L) \leq \nu - 1$, (see Elliot & Wolovich [11]). Consequently, the number of parameters to identify, $N_\theta = p(p+m)\nu$. Note that $\{N_L, D_L\}$ is generally not unique.

2.4 Parameter estimation algorithm

We will assume that the following normalized least-squares identification algorithm with covariance resetting is used:

$$\dot{\phi} \;=\; \dot{\theta} \;= -g\frac{P\psi e_2^T}{1 + \gamma\psi^T\psi} \quad \text{with} \quad g, \gamma > 0$$

$$\dot{P} \;=\; -g\frac{P\psi\psi^T P}{1+\gamma\psi^T\psi} \quad \text{with} \quad P(0) = P(\tau_k) = k_0 I > 0 \tag{2.29}$$

where e_2 is the identifier error (2.15) and $\{\tau_k\}$ are the resetting time instants such that

$$\{\tau_k\} = \{\tau_k | \lambda_{\max}(P(\tau_k^- - \delta_{\tau_k})) = k_1 \quad \text{with} \quad 0 \le k_1 < k_0 \quad \text{and} \quad 0 \le \delta_{\tau_k} \le k_2\}$$

where $\lambda_{\max}$ denotes the largest eigenvalue. The matrix P is discontinuous at the resetting instants τ_k, with the limit on the right $P(\tau_k)$ equal to a predetermined matrix $k_0 I$. A degree of freedom has been left in the definition of the resetting instants. Resetting occurs within a fixed period of time k_2 after $\lambda_{\max}(P(t))$ has reached a certain minimum level k_1, but the specific time instant is left free at this stage. The time delays δ_{τ_k} are introduced to account for the fact that $\lambda_{\max}(P(t))$ will only be computed at discrete time intervals. If $k_1 = 0$, the standard least-squares algorithm without resetting is obtained. For $k_1 > 0$, the algorithm will reset whenever the P matrix becomes too small.

The following definitions are needed before stating the properties of the estimation algorithm.

Definition 2.1 : Persistent Excitation
A bounded vector ψ is *persistently exciting* (PE) iff

$$\exists\, \alpha > 0 \;\; \exists\, \delta > 0 \quad \text{such that} \quad \int_{t_0}^{t_0+\delta} \psi\psi^T \, d\tau \;\ge\; \alpha I \quad \forall\, t_0 \ge 0$$

Definition 2.2 : Sufficient Excitation
A bounded vector ψ is *sufficiently exciting* (SE) iff

$$\exists\, \alpha > 0 \;\text{ such that }\; \forall\, t_0 \ge 0 \;\; \exists\, \delta(t_0) > 0 \quad \text{such that} \quad \int_{t_0}^{t_0+\delta} \psi\psi^T \, d\tau \;\ge\; \alpha I$$

Note that if a signal is PE then it is also SE, but the converse is not true.

Lemma 2.3 : Properties of the estimation algorithm
Assuming that $\psi \in L_{\infty e}$, the estimation algorithm (2.29) has the following properties

1. $0 \le P \le k_0 I$, $\quad k_0^{-1} I \le P^{-1} \in L_{\infty e}$, $\quad$ and $\quad \frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}}$, $\dot{P}$, $\pi = \frac{P\psi}{1+\|\psi_t\|_\infty} \in L_\infty$.

2. $(\tau_{k+1} - \tau_k)$ is bounded below $\forall\, k$.

3. ϕ, $\frac{e_2}{(1+\gamma\psi^T\psi)^{1/2}}$, $\dot{\phi}$, $\beta = \frac{\phi^T\psi}{1+\|\psi_t\|_\infty} \in L_\infty$.

4. $\frac{e_2}{(1+\gamma\psi^T\psi)^{1/2}}$, $\dot{\phi}$, $\beta \in L_2$.

5. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE or if $k_1 = 0$ (no resettings):

   - $\{\tau_k\}$ is a finite set.
   - $\frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}}$, $\dot{P}$, $\pi = \frac{P\psi}{1+\|\psi_t\|_\infty} \in L_2$.
   - $\dot{\phi}$, $\dot{P} \in L_1$.
   - $P(t)$ and $\phi(t)$ converge to some P_∞ and ϕ_∞.

6. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE:

   - $\lim_{t\to\infty} \phi(t) = 0$.

7. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is PE and $k_1 > 0$:

   - $\lim_{t\to\infty} \phi(t) = 0$ exponentially.

8. $|\phi_i(t)| \leq |\phi_i(0)|$, $\forall\, t \geq 0$, where ϕ_i is the i-th column of ϕ.

The proof is in the appendix.

Comments:

The specific adaptation algorithm is chosen for its special properties given in Lemma 2.3. Most important is the fact that the parameter error ϕ converges (although not necessarily to zero) independent of richness properties of ψ. It may converge to zero or to some other value. Also, if ψ is not SE or if $k_1 = 0$ (no resettings) $P\psi$ will have the same properties as e_2 and there will be a finite number of resettings. These properties result from the use of $\lambda_{\max}(P(t))$ for the resetting instead of $\lambda_{\min}(P(t))$ as in Sastry & Bodson [3]. When there is not sufficient excitation, the algorithm will not reset the covariance after some time and the algorithm has the same properties as when covariance resetting is not used. Finally, it should be noted that the exponential convergence under PE conditions is guaranteed (rather than asymptotic convergence) when the covariance resetting is used. However, the stability results will be valid whether the resetting is used or not and whether there is excitation or not.

2.5 Stability

2.5.1 Stability of the direct MRAC algorithm

Stability proofs for MIMO systems follow similar paths as for SISO systems. Difficulties arise, mainly because of singularities that are encountered, such as when the matrix C_0 is singular in the MRAC algorithm. These problems are also present in the SISO case, but become more complex in MIMO systems.

Consider the problem of singularity of C_0 in the MRAC algorithm. Attempts to prove the global stability of the algorithm reveal that the inverse of the matrix used by the controller must be bounded. This boundedness is not guaranteed by the adaptation algorithm described in the previous section so that the adaptive control scheme must be modified when the estimate of C_0 comes too close to singularity. There has been a significant research effort to resolve this problem in the SISO case and several solutions have been proposed. Their extension to MIMO systems is far from trivial however. A New MIMO parameter transformation is presented in Chapter 4 which allows the controller parameter of a MRAC scheme to stay away from singularity. In this chapter, however, we will rely on stability results that are based on rather restrictive assumptions, but are easily derived from the current literature.

Before stating any stability results, the following definitions are needed. Since the regressor vector ψ is made of filtered input and output signals, define the transfer function matrix, $H_{\psi u}(s)$, as the transfer function between ψ and u. We also define the *model signals* ψ_m as the signals ψ when the parameter error $\phi = 0$ (model matching is achieved)

$$\psi_m = H_{\psi u} P^{-1} H[r] = H_{\psi_m r}[r] \tag{2.30}$$

It can be verified that $H_{\psi_m r}(s)$ is stable and strictly proper, since $P(s)$ is minimum phase. Finally, define the *regressor error* as

$$e_\psi = \psi - \psi_m$$

and the *output error* as

$$e_0 = y_p - y_m$$

Given these definitions, the following stability results can be proved:

Theorem 2.1 : Global stability of the direct MRAC algorithm - K_p known

Consider the MIMO MRAC system described in Section 2.3.1. Assume that the high-frequency gain matrix $K_p = C_0^{*-1}$ is known, so that the control parameter matrix C_0 does not need to be estimated. Assume that the recursive least-squares algorithm (2.29) is used for parameter identification. If the reference input $r \in L_\infty$ and is piecewise continuous, then

1. all the states of the adaptive control system are bounded.
2. the regressor error $e_\psi \in L_2 \cap L_\infty$ and $\lim_{t\to\infty} e_\psi = 0$.
3. the output error $e_0 \in L_\infty$ and $\lim_{t\to\infty} e_0 = 0$.

The proof of this theorem is a relatively straightforward MIMO extension of the SISO proof in Sastry & Bodson [3] and is omitted here. Requiring the *a priori* knowledge of the high-frequency gain matrix K_p is not very practical however. A slightly less restrictive solution is proposed in Singh & Narendra [9] and Tao & Ioannou [8], which assume that a matrix S such that $SK_p + (SK_p)^T > 0$ is known. In the SISO case, this is equivalent to assuming the knowledge of the sign of the high-frequency gain. Assuming a less restrictive type of prior knowledge, the following stability result can also be proved:

Theorem 2.2 : Local stability of the direct MRAC algorithm - K_p unknown
Consider the MIMO MRAC system described in Section 2.3.1. Assume that the columns of the initial parameter error, $\phi_i(0)$, are sufficiently small that, for all vectors ϕ_i such that $\|\phi_i\| \leq \|\phi_i(0)\|$, C_0 is nonsingular. Assume that the recursive least-squares algorithm (2.29) is used for parameter identification. If the reference input $r \in L_\infty$ and is piecewise continuous, then

1. all the states of the adaptive control system are bounded.
2. the regressor error $e_\psi \in L_2 \cap L_\infty$ and $\lim_{t\to\infty} e_\psi = 0$.
3. the output error $e_0 \in L_\infty$ and $\lim_{t\to\infty} e_0 = 0$.

Given the convergence properties of the least-squares estimation algorithm (see Lemma 2.3), $\|\phi_i(t)\| \leq \|\phi_i(0)\|\ \forall\, t > 0$. Therefore, if the initial parameter error is "sufficiently small", $C_0(t)$ will remain nonsingular $\forall\, t > 0$. Again, the proof is a direct extension of the proof in the SISO case and is omitted here.

The condition on $\phi(0)$ requires a strong *a priori* information. This condition is necessary to avoid singularity regions. It will be shown in Chapter 4 how this condition can be relaxed.

2.5.2 Stability of the direct PPAC algorithm

Similarly to the MRAC algorithm, the regressor vector is made of filtered input and output signals, so that a transfer function matrix $H_{\psi u}(s)$ between ψ and u can be defined. Similarly, the model signals ψ_m are defined as the signals ψ when $\phi = 0$

$$\psi_m = H_{\psi u} D_R D_M^{-1}[r] = H_{\psi_m r}[r] \tag{2.31}$$

A proof of the global stability and convergence of the error signals e_0 and e_ψ to zero is currently not available without input conditions for the PPAC algorithm. At the price of some additional prior information, the following result can be proved:

Theorem 2.3 : Local stability of the direct PPAC algorithm
Consider the MIMO PPAC system described in Section 2.3.3. Assume that the columns of the initial parameter error, $\phi_i(0)$, are sufficiently small that, for all vectors ϕ_i such that $\|\phi_i\| \leq \|\phi_i(0)\|$, the closed-loop poles of the system are stable. Assume that the recursive least-squares algorithm (2.29) is used for parameter identification. If the reference input $r \in L_\infty$ and is piecewise continuous, then

1. all the states of the adaptive control system are bounded.
2. the regressor error $e_\psi \in L_\infty$.
3. the output error $e_0 \in L_\infty$.

The proof is in the appendix.

Again, the condition on $\phi(0)$ requires a strong *a priori* information. This condition is necessary to avoid singularity regions. Unfortunately, it is still very much an open problem. It might be noted that some stronger results are available. In the SISO case, it is proved in Elliot *et al.* [62] that, if the regressor vector ψ is *persistently exciting (PE)*, then all states remain bounded. Furthermore, e_0 and e_ψ will converge to zero. However, the proof assumes that the controller parameters are updated at discrete time instants. A similar result is obtained by Willner *et al.* [34] for a discrete time MIMO algorithm. For a discrete time SISO algorithm using a different setting, it is showed in Kamen [63] that, if the input is *sufficiently exciting*, then the parameter estimates θ will converge "sufficiently close" to their true values θ^*, so that all states remain bounded. Finally, complex modifications are proposed in Janecki [64] to guarantee global stability without input conditions for a discrete time SISO algorithm. Unfortunately, an auxiliary estimation algorithm must be introduced for the implementation.

2.6 Comparison

This section makes a general comparison of the algorithms presented in this chapter in terms of necessary *a priori* information and number of parameters to identify. Table 2.1 summarizes the previous results concerning the *a priori* information and the number of adaptive parameters required by the different algorithms. Only the *a priori* information relevant to the *parameterization* is shown in Table 2.1. In other words, this is the information needed to defined the structure of the controller, including the number of its parameters. Other assumptions may be needed for stability considerations. For example, a MRAC algorithm requires minimum phase assumptions that are not needed for PPAC. Other assumptions, such as those necessary to guarantee that C_0^{-1} is bounded may also be needed. These assumptions are summarized in Table 2.2. Note that a complete and correct proof of stability for the extended MRAC algorithm does not exists in the literature. Therefore, at this stage, only local stability will be assumed.

Algorithm	*A priori* information	N_θ
Direct MRAC	$H(s), \nu \geq \nu_{\max}$	$2p^2\nu$
Extended MRAC	$\{r_i\}$, $\nu \geq \nu_{\max}$	$\geq 2p^2\nu$
Direct PPAC	$\{\mu_i\}, \nu \geq \nu_{\max}$	$\leq 2m^2(\nu-1) + 2mp\nu + m(m\mu_{\max} - n)$ $\geq 2m^2(\nu-1) + 2mp\nu$
Recursive identification	$\nu \geq \nu_{\max}$	$p(p+m)\nu$

Table 2.1: *A priori* information and number of parameters

Algorithm	*A priori* information	Stability result
Direct MRAC	Minimum phase, $\phi(0)$ sufficiently small	Local stability
Direct MRAC	Minimum phase, K_p	Global stability
Extended MRAC	Minimum phase, $\phi(0)$ sufficiently small	Local stability
Direct PPAC	$\phi(0)$ sufficiently small	Local stability

Table 2.2: *A priori* information and stability

Chapter 3

Parameter convergence

3.0 Introduction

Several MIMO MRAC algorithms have been proposed in the literature, *e.g.*, Goodwin & Long [20], Elliot & Wolovich [6, 11], Singh & Narendra [9, 10], Dugard *et al.* [12], Johansson [13], Ortega *et al.* [14], Das [7], Dion *et al.* [50], and Tao & Ioannou [8]. The direct pole placement approach was investigated by Elliot *et al.* [32], Djaferis *et al.* [33], and Elliot & Wolovich [11].

The emphasis of these papers has been on finding parameterizations for the controllers, (a problem much more complex for multivariable systems than for SISO systems), proving stability (in a few cases), and trying to extend the schemes in ways to reduce the amount of *a priori* information required. Close inspection reveals that parameter convergence is not guaranteed for these schemes, even with rich input signals, because they rely on parameterizations that do not define the parameters uniquely (in the previous chapter (Chapter 2), the controller parameters were not uniquely defined). A parameterization that defines the controller parameters uniquely will be called *identifiable*. The parameterizations presented in the previous chapter (Chapter 2) were not identifiable, since θ^* was not uniquely defined. When the parameterization is not identifiable, the estimate of the parameters θ cannot converge to a unique value. Instead, θ converges into a set. Therefore, parameter convergence cannot be guaranteed, even with rich input signals.

In this chapter, we show how to modify current continuous-time direct MRAC and PPAC algorithms and a recursive identifier, to guarantee identifiability. In other words, we present direct MRAC and PPAC algorithms and a recursive identifier, using parameterizations for which the parameters are uniquely defined. We extend the work of Willner *et al.* [34] which considered the problem for a discrete-time adaptive pole placement algorithm.

Parameter convergence is important for robustness to noise and unmodeled dynamics (*cf.* Sastry & Bodson [3]). At the end of this chapter, an example illustrates this fact by showing two MRAC algorithms (one with an identifiable parameterization and the other with a non-identifiable parameterization) applied to a system with noise and unmodeled dynamics. Only

the scheme based on the identifiable parameterization remained stable.

In this chapter, we also derive simple frequency domain conditions on the inputs that guarantee parameter convergence to the nominal values for the adaptive control schemes presented in the previous chapter (Chapter 2). Similar conditions were obtained, using generalized harmonic analysis, in multivariable recursive identification by de Mathelin & Bodson [65], and in SISO adaptive control by Boyd & Sastry [66]. However, the conditions in multivariable adaptive control are not trivial extensions of the conditions in the SISO case. An important difference with the SISO case is that the necessary and the sufficient conditions for parameter convergence may be different. Indeed, parameter convergence may depend on the location of the spectral components of the inputs. For the same number of frequencies in the inputs, convergence may depend on the locations of these frequencies. This is not the case with SISO systems where only the number of frequencies is relevant (*cf.* Boyd & Sastry [66]). Note that a summary of these results was presented in de Mathelin & Bodson [67].

On the basis of the results, the adaptive schemes and a recursive identification scheme are compared with respect to the requirements for parameter convergence. We note that although the parameterizations under consideration are all identifiable, they have different numbers of parameters and require different conditions for parameter convergence. We present specific examples illustrating these differences and show cases where an indirect scheme converges, although the direct scheme does not, and *vice versa*. This shows that the so-called equivalence between direct and indirect schemes is not as obvious as it appears from the SISO case. Depending on the application under consideration, the designer may find advantages to direct *vs* indirect approaches. In some cases, the computations required by a direct scheme may be significantly less than for an indirect scheme. The reverse may also be true and a decision must be made on a case by case basis.

3.1 Preliminaries

We will denote by $\rho_{\max}$ the maximum of the pseudo-observability indices $\{\rho_i\}$ (*cf.* Definition 1.6). Similarly, $\kappa_{\max}$ will denote the maximum of the pseudo-controllability indices $\{\kappa_i\}$ (*cf.* Definition 1.7).

The following lemma will be useful to to define the identifiable parameterizations.

Lemma 3.1 : Null matrix condition

Given $\{N_L(s), D_L(s)\}$ and $\{N_R(s), D_R(s)\}$, a left and a right MFD of a $(p \times m)$ strictly causal system with observability indices $\{\nu_i\}$, pseudo-observability indices $\{\rho_i\}$, controllability indices $\{\mu_i\}$, and pseudo-controllability indices $\{\kappa_i\}$,
if $D(s) \in R^{q\times p}[s]$, $N(s) \in \Re^{q\times m}[s]$, with q arbitrary, are such that:

$$D(s)N_R(s) = N(s)D_R(s) \ \forall\, s \quad \text{and} \quad \partial c_i D \leq \nu_i - 1 \quad \text{or} \quad \partial c_i D \leq \rho_i - 1 \quad \forall i$$

or if $D(s) \in R^{m\times q}[s]$, $N(s) \in \Re^{p\times q}[s]$, with q arbitrary, are such that:

$$N_L(s)D(s) = D_L(s)N(s) \;\; \forall\, s \quad \text{and} \quad \partial r_i D \le \mu_i - 1 \quad \text{or} \quad \partial r_i D \le \kappa_i - 1 \;\; \forall i$$

then

$$D(s) = 0 \qquad N(s) = 0$$

The proof of this lemma is in the appendix.

3.2 Identifiable parameterization

The same controller structure as in Section 2.2.1 is used.

3.2.1 Model reference adaptive control

Given assumptions (A1), (A2), (A3) of Section 2.2.2, we show that there exist unique controller parameters values such that y_p matches y_m.

Proposition 3.1 : Unique model reference matching equality
If $\partial\lambda_i = \nu_{\max} - 1$, $\exists$ *unique* controller parameters $C_0^* \in \Re^{p\times p}$, $C^*(s)$, $D^*(s) \in \Re^{p\times p}[s]$, solution of the Diophantine equation:

$$N_R[(\Lambda - C^*)D_R - D^* N_R]^{-1}\Lambda C_0^* = H \tag{3.1}$$

such that model matching is achieved (the transfer function from r to y_p is $H(s)$), $\Lambda^{-1}D^*$ is proper, $\Lambda^{-1}C^*$ is strictly proper, and D^* satisfies the following constraint

$$\partial c_i D^* \le \nu_i - 1 \quad \forall i \tag{3.2}$$

In particular, $C_0^* = K_p^{-1}$ nonsingular and $\partial r_i C^* < \partial\lambda_i$. Similarly, $\exists$ unique $\bar{C}^*(s)$ and $\bar{D}^*(s)$ $\in \Re^{p\times p}[s]$ such that model matching is achieved, $\Lambda^{-1}\bar{D}^*$ and $\Lambda^{-1}\bar{C}^*$ are proper, $\bar{C}^*_{0\infty} = I - K_p$, and $\partial c_i \bar{D}^* \le \nu_i - 1$.

Proof
The proof follows the same lines than the proof of Proposition 2.1. The matching equality (3.1) is equivalent to

$$C_0^* H^{-1} P = I - \Lambda^{-1}C^* - \Lambda^{-1}D^* P \tag{3.3}$$

Now, let $\{N_L, D_L\}$ be the canonical left MFD of $P(s)$ (D_L column reduced, $\partial c_i D_L(s) = \nu_i$, *cf.* Definition 1.13). Using the polynomial matrix division (Lemma 2.2), divide $\Lambda K_p^{-1} H^{-1}$ on the right by D_L, then $\exists$ $Q(s)$, $R(s) \in \Re^{p\times p}[s]$ such that

$$\Lambda K_p^{-1} H^{-1} = Q D_L + R \quad \text{and} \quad R D_L^{-1} \text{ is strictly proper}$$

and $\partial c_i R < \partial c_i D_L = \nu_i$. Then, let

$$D^* = -R = QD_L - \Lambda K_p^{-1} H^{-1} \qquad C^* = \Lambda - QN_L \qquad C_0^* = K_p^{-1}$$

It can be easily verified that the given C_0^*, C^*, D^* solve the matching equality (3.3) with $\partial c_i D^* \leq \nu_i - 1$. Furthermore, since $\partial\lambda_i = \nu_{\max} - 1$, $\Lambda^{-1} D^*$ is proper. On the other hand,

$$\lim_{s\to\infty} \Lambda^{-1} C^* = \lim_{s\to\infty} (I - K_p^{-1} H^{-1} P - \Lambda^{-1} D^* P) = I - I = 0$$

so that $\Lambda^{-1} C^*$ is strictly proper and $\partial r_i C^* \leq \partial\lambda_i - 1 = \nu_{\max} - 2$.

Now, suppose that $C_0^* + \Delta C_0$, $C^* + \Delta C$, $D^* + \Delta D$ is another solution to the matching equality. Then

$$\Delta C_0 H^{-1} P + \Lambda^{-1} \Delta C + \Lambda^{-1} \Delta D P = 0$$

and

$$\lim_{s\to\infty} \Delta C_0 H^{-1} P = \Delta C_0 K_p = 0 \quad \Rightarrow \quad \Delta C_0 = 0$$

$$\Rightarrow \quad \Delta C D_R + \Delta D N_R = 0 \quad \text{and} \quad \partial c_i \Delta D \leq \nu_i - 1$$

by the null matrix condition (Lemma 3.1)

$$\Delta C = \Delta D = 0$$

So that the solution is unique.

Similarly, dividing ΛH^{-1} on the right by D_L, then $\exists\ Q(s),\ R(s) \in \Re^{p\times p}[s]$ such that

$$\Lambda H^{-1} = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

Let

$$\bar{D}^* = \bar{Q} D_L - \Lambda H^{-1} \qquad \bar{C}^* = \Lambda - \bar{Q} N_L$$

then $\partial c_i \bar{D}^* \leq \nu_i - 1$, $\partial r_i \bar{C}^* \leq \partial\lambda_i$, $\bar{C}^*_{0_\infty} = I - K_p$, and the given $\bar{C}^*$ and $\bar{D}^*$ are a solution of the matching equality

$$N_R[(\Lambda - \bar{C}^*) D_R - \bar{D}^* N_R]^{-1} \Lambda = H$$

Furthermore, this solution is unique. $\square$

Corollary 3.1

If $\partial\lambda_i \geq \rho_{\max} - 1$, $\exists$ unique matrices $C_0^* \in \Re^{p\times p}$, $C^*(s)$, $D^*(s) \in \Re^{p\times p}[s]$, such that model matching is achieved, $\Lambda^{-1} D^*$ is proper, $\Lambda^{-1} C^*$ is strictly proper, $C_0^* = K_p^{-1}$, and $\partial c_i D^* \leq \rho_i - 1$. Similarly, $\exists$ unique matrices $\bar{C}^*(s)$ and $\bar{D}^*(s) \in \Re^{p\times p}[s]$ such that model matching is achieved, $\Lambda^{-1}\bar{D}^*$ and $\Lambda^{-1}\bar{C}^*$ are proper, $\bar{C}^*_{0_\infty} = I - K_p$. and $\partial c_i \bar{D}^* \leq \rho_i - 1$.

The proof is similar to the proof of Proposition 3.1 with the pseudo-canonical left MFD replacing the canonical left MFD.

To guarantee parameter convergence of the MRAC scheme, the uniqueness of θ^* must be guaranteed (θ^* must be identifiable). From Proposition 3.1, the observability indices $\{\nu_i\}$ must be known and θ^* must be constrained so that $\partial c_i D^* \leq \nu_i - 1$ and $\nu = \nu_{\max}$. This can be done simply by having the j-th column of $D_i = 0$ if $\nu_j < i$. An advantage of uniqueness is that the number of parameters N_θ becomes smaller, $N_\theta = p^2\nu_{\max} + pn$. Note also that $N_\theta = 2pn$ if the observability indices are all equal and that $N_\theta = 2n$ in the SISO case, as expected.

From Corollary 3.1, uniqueness can also be achieved if a set of pseudo-observability indices $\{\rho_i\}$ is known, $\partial\lambda_i = \rho - 1 = \rho_{\max} - 1$, and θ^* is constrained so that $\partial C^* \leq \rho_{\max} - 2$ and $\partial c_i D^* \leq \rho_i - 1$. Then, $N_\theta = p^2\rho_{\max} + pn$.

In the case of the extended MRAC (Section 2.3.2), the uniqueness of $(\Sigma^*, \bar{C}^*, \bar{D}^*)$ cannot be guaranteed even with the constraint $\partial c_i \bar{D}^* \leq \nu_i - 1$, ($\bar{C}^*$ and $\bar{D}^*$ are unique only if Σ^* is known). Therefore, parameter convergence cannot be ensured.

3.2.2 Adaptive pole placement

Given assumptions (A1), (A2), (A3) of Section 2.2.3, we show that there exist unique controller parameters values such that y_p matches y_m.

Proposition 3.2 : Unique pole placement matching equality

If $\partial\lambda_i = \nu - 1$, and $\nu \geq \nu_{\max}$, $\exists C_0^* \in \Re^{m\times m}$, $C^*(s) \in \Re^{m\times m}[s]$, $D^*(s) \in \Re^{m\times p}[s]$, solution of the Diophantine equation:

$$[(\Lambda - C^*)D_R - D^*N_R]^{-1}\Lambda C_0^* = D_M^{-1} \tag{3.4}$$

such that model matching is achieved, $\Lambda^{-1}D^*$ is proper, $\Lambda^{-1}C^*$ is strictly proper, and D^* satisfies the following constraint

$$\partial c_i D^* \leq \nu_i - 1 \quad \forall i \tag{3.5}$$

In particular, $C_0^* = \Gamma_c[D_R]$ nonsingular, and $\partial r_i C^* < \partial\lambda_i$. Similarly, $\exists\ \bar{C}^* \in \Re^{m\times m}[s]$ and $\bar{D}^* \in \Re^{m\times p}[s]$ such that model matching is achieved, $\Lambda^{-1}\bar{D}^*$ and $\Lambda^{-1}\bar{C}^*$ are proper, $\bar{C}^*_{0_\infty} = I - \Gamma_c[D_R]^{-1}$, and $\partial c_i \bar{D}^* \leq \nu_i - 1$.

Proof

The proof is similar to the model reference case. The matching equality (3.4) is equivalent to

$$C_0^* D_M D_R^{-1} = I - \Lambda^{-1}C^* - \Lambda^{-1}D^*P \tag{3.6}$$

Since $\{N_R, D_R\}$ are coprime, by the Bezout identity (Lemma 2.1), $\exists\ U_L^* \in \Re^{m\times p}[s]$, $V_L^* \in \Re^{m\times m}[s]$ such that

$$U_L^* N_R + V_L^* D_R = I \quad \Leftrightarrow \quad U_L^* P + V_L^* = D_R^{-1}$$

Now, let $\{N_L, D_L\}$ be the canonical left MFD of $P(s)$. Using the polynomial matrix division (Lemma 2.2), divide $\Lambda\Gamma_c[D_R]D_M U_L^*$ on the right by D_L, then $\exists\ Q(s),\ R(s) \in \Re^{m\times p}[s]$ such that

$$\Lambda\Gamma_c[D_R]D_M U_L^* = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

and $\partial c_i R < \partial c_i D_L = \nu_i$. Then, let

$$D^* = -R = QD_L - \Lambda\Gamma_c[D_R]D_M U_L^* \qquad C^* = \Lambda - QN_L - \Lambda\Gamma_c[D_R]D_M V_L^* \qquad C_0^* = \Gamma_c[D_R]$$

It is easy to verify that the given C_0^*, C^*, D^* solve the matching equality (3.6) and $\partial c_i D^* \leq \nu_i - 1$. Furthermore, since $\partial\lambda_i \geq \nu - 1$, $\Lambda^{-1}D^*$ is proper. On the other hand

$$\lim_{s\to\infty} \Lambda^{-1}C^* = \lim_{s\to\infty}(I - \Gamma_c[D_R]D_M(U_L^*P + V_L^*) - \Lambda^{-1}D^*P) = I - I = 0$$

so that $\Lambda^{-1}C^*$ is strictly proper and $\partial r_i C^* \leq \partial\lambda_i - 1$.

Now, suppose that $C_0^* + \Delta C_0$, $C^* + \Delta C$, $D^* + \Delta D$ is another solution to the matching equality. Then

$$\Delta C_0 D_M D_R^{-1} + \Lambda^{-1}\Delta C + \Lambda^{-1}\Delta DP = 0$$

and

$$\lim_{s\to\infty} \Delta C_0 D_M D_R^{-1} = \Delta C_0 \Gamma_c[D_R]^{-1} = 0 \quad \Rightarrow \quad \Delta C_0 = 0$$

$$\Rightarrow \quad \Delta C D_R + \Delta D N_R = 0 \quad \text{and} \quad \partial c_i \Delta D \leq \nu_i - 1$$

by the null matrix condition (Lemma 3.1)

$$\Rightarrow \quad \Delta C = \Delta D = 0$$

So that the solution is unique.

Similarly, dividing $\Lambda D_M U_L^*$ on the right by D_L, then $\exists\ Q(s),\ R(s) \in \Re^{m\times p}[s]$ such that

$$\Lambda D_M U_L^* = QD_L + R \quad \text{and} \quad RD_L^{-1} \text{ is strictly proper}$$

Let

$$\bar{D}^* = \bar{Q}D_L - \Lambda D_M U_L^* \qquad \bar{C}^* = \Lambda - \bar{Q}N_L - \Lambda D_M V_L^*$$

then $\partial c_i \bar{D}^* \leq \nu_i - 1$, $\partial r_i \bar{C}^* \leq \partial\lambda_i$, $\bar{C}_{0_\infty}^* = I - \Gamma_c[D_R]^{-1}$, and the given $\bar{C}^*$ and $\bar{D}^*$ are a solution of the matching equality

$$[(\Lambda - \bar{C}^*)D_R - \bar{D}^*N_R]^{-1}\Lambda = D_M^{-1}$$

Furthermore, this solution is unique. □

Corollary 3.2

If $\partial\lambda_i = \rho - 1$ and $\rho \geq \rho_{\max}$, $\exists$ unique matrices $C_0^* \in \Re^{m\times m}$, $C^*(s) \in \Re^{m\times m}[s]$, $D^*(s) \in \Re^{m\times p}[s]$, such that model matching is achieved, $\Lambda^{-1}D^*$ is proper, $\Lambda^{-1}C^*$ is strictly proper, $C_0^* = \Gamma_c[D_R]$ and $\partial c_i D^* \leq \rho_i - 1$. Similarly, $\exists$ unique matrices $\bar{C}^* \in \Re^{m\times m}[s]$ and $\bar{D}^* \in \Re^{m\times p}[s]$ such that model matching is achieved, $\Lambda^{-1}\bar{D}^*$ and $\Lambda^{-1}\bar{C}^*$ are proper, $\bar{C}_{0_\infty}^* = I - \Gamma_c[D_R]^{-1}$, and $\partial c_i \bar{D}^* \leq \rho_i - 1$.

The proof is similar to the proof of Proposition 3.2 with the pseudo-canonical left MFD replacing the canonical left MFD.

Now, we show that there exist unique matrices U_L^* and V_L^* satisfaying the Bezout identity (2.20).

Proposition 3.3 : Uniqueness of the Bezout identity

Given a right coprime MFD $\{N_R, D_R\}$, $\exists$ unique matrices $U_L^* \in \Re^{m\times p}[s]$, $V_L^* \in \Re^{m\times m}[s]$ such that

$$V_L^* D_R + U_L^* N_R = I \quad \text{and} \quad \partial c_i U_L^* \leq \nu_i - 1$$

Furthermore, $\partial V_L^* < \partial U_L^* \leq \nu_{\max} - 1$.

Proof

Since $\{N_R, D_R\}$ are coprime, we know there exists a solution $\bar{U}_L \in \Re^{m\times p}[s]$, $\bar{V}_L \in \Re^{m\times m}[s]$ such that

$$\bar{V}_L D_R + \bar{U}_L N_R = I$$

Divide $\bar{U}_L$ on the right by D_L, where $\{N_L, D_L\}$ is a canonical left MFD (see Definition 1.13), then by the polynomial matrix division (Lemma 2.2) $\exists$ Q_L and $U_L^* \in \Re^{m\times p}[s]$ such that

$$\bar{U}_L = Q_L D_L + U_L^* \quad \text{and} \quad U_L^* D_L^{-1} \text{ is strictly proper}$$

and $\partial c_i U_L^* \leq \nu_i - 1$. Now let

$$V_L^* = \bar{V}_L + Q_L N_L$$

It can be easily verified that U_L^*, V_L^* is also a solution of the Bezout identity. Now, suppose that $U_L^* + \Delta U_L$, $V_L^* + \Delta V_L$ is another solution such that $\partial c_i(U_L^* + \Delta U_L) \leq \nu_i - 1$. Then

$$\Delta U_L N_R + \Delta V_L D_R = 0 \quad \text{and} \quad \partial c_i \Delta U_L \leq \nu_i - 1$$

by the null matrix condition (Lemma 3.1)

$$\Rightarrow \quad \Delta U_L = \Delta V_L = 0$$

so that the solution is unique.

Furthermore, from (2.20), we have that $\partial V_L^* < \partial U_L^* \leq \nu_{\max} - 1$. $\square$

Corollary 3.3

Given a right coprime MFD $\{N_R, D_R\}$, $\exists$ unique matrices $U_L^* \in \Re^{m\times p}[s]$, $V_L^* \in \Re^{m\times m}[s]$ such that

$$V_L^* D_R + U_L^* N_R = I \quad \text{and} \quad \partial c_i U_L^* \leq \rho_i - 1 \quad \forall\, i$$

Furthermore, $\partial V_L^* < \partial U_L^* \leq \rho_{\max} - 1$.

The proof is similar to the proof of Proposition 3.3 with the pseudo-canonical left MFD replacing the canonical left MFD.

Similarly, to Section 2.3.3, we must guarantee that, for all inputs u, equation (2.21) has a solution which is also the solution of the pole placement matching equality (2.19) and the Bezout identity (2.20).

Proposition 3.4 : Unique equivalence of (2.21) with (2.19) and (2.20)

If

$$\partial\lambda_i = \mu_{\max} - \mu_i + \nu_{\max} - 1$$

then there exists a solution $\bar{C}^*, \bar{D}^*$ of the matching equality (2.19) such that

$$\begin{aligned} \partial c_i \bar{D}^* &\leq \nu_i - 1 \\ \partial r_i \bar{C}^* &\leq \mu_{\max} - \mu_i + \nu_{\max} - 1 \\ \partial c_i \bar{C}^* &\leq \mu_{\max} - \mu_i + \nu_{\max} - 2 \\ \bar{C}^*_{\infty_{ij}} &= (I - \Gamma_c[D_R]^{-1})_{ij} = 0 \quad \text{if} \quad i \geq j \end{aligned} \tag{3.7}$$

Furthermore, under these constraints, any solution of (2.21) valid for all u is also a solution of (2.19) and (2.20).

The proof is almost identical to the proof of Proposition 2.4 and will be omitted here.

Corollary 3.4

If

$$\partial\lambda_i = \mu_{\max} - \mu_i + \rho_{\max} - 1$$

then there exists a solution $\bar{C}^*, \bar{D}^*$ of the matching equality (2.19) such that

$$\begin{aligned} \partial c_i \bar{D}^* &\leq \rho_i - 1 \\ \partial r_i \bar{C}^* &\leq \mu_{\max} - \mu_i + \rho_{\max} - 1 \\ \partial c_i \bar{C}^* &\leq \mu_{\max} - \mu_i + \rho_{\max} - 2 \\ \bar{C}^*_{\infty_{ij}} &= (I - \Gamma_c[D_R]^{-1})_{ij} = 0 \quad \text{if} \quad i \geq j \end{aligned} \tag{3.8}$$

Furthermore, under these constraints, any solution of (2.21) valid for all u is also a solution of (2.19) and (2.20).

The proof is similar to the proof of Proposition 2.4 but relies on Corollary 3.2 instead of Proposition 2.2.

Consequently, uniqueness of θ^* is not guaranteed, unless the observability indices $\{\nu_i\}$ are known (*cf.* Propositions 3.2, 3.3, and 3.4), and θ^* is constrained so that $\nu = \nu_{\max}$, $\partial c_i D^* \leq \nu_i - 1$, and $\partial c_i U_L^* \leq \nu_i - 1$. Then, N_θ is smaller

$$2m^2(\nu_{\max} - 1) + 2mn \leq N_\theta \leq 2m^2(\nu_{\max} - 1) + m(m\mu_{\max} + n)$$

Note that $N_\theta = 2m^2(\nu_{\max}-1)+2mn$ if the controllability indices are all equal, and $N_\theta = 4n-2$ in the SISO case. Our algorithm is comparable to the discrete-time algorithm of Willner *et al.* [34], with some nontrivial differences between the continuous-time and discrete-time algorithms.

From Corollaries 3.2, 3.3, and 3.4 uniqueness can also be achieved if a set of pseudo-observability indices $\{\rho_i\}$ is known, $\rho = \rho_{\max}$, $\partial l = \rho - 1$, $\partial \lambda_i = \mu_{\max} - \mu_i + \rho - 1$, and θ^* is constrained so that $\partial V_L^* \leq \rho_{\max} - 2$, $\partial c_i U_L^* \leq \rho_i - 1$, C^* obeys to (3.8), and $\partial c_i D^* \leq \rho_i - 1$. Then,

$$2m^2(\rho_{\max}-1)+2mn \;\leq\; N_\theta \;\leq\; 2m^2(\rho_{\max}-1)+m(m\mu_{\max}+n)$$

3.2.3 Recursive identification

To guarantee uniqueness of θ^*, a pseudo-canonical or a canonical parameterization must be used for the recursive identification (*cf.* Section 2.3.4).

Identification of pseudo-canonical MFD

Assume that $\{N_L(s), D_L(s)\}$ is a pseudo-canonical left MFD of the strictly proper transfer function matrix P(s), with pseudo-observability indices $\{\rho_i\}$. Define $\Lambda(s) = \mathrm{diag}\{(s+a)^{\rho_{\max}}\}$ and $L(s) = \mathrm{diag}\{(s+a)^{\rho_{\max}-\rho_i}\}$ with $a > 0$, then

$$L^{-1}[y_p] = \Lambda^{-1}N_L[u] - \Lambda^{-1}(D_L - L^{-1}\Lambda)[y_p] \tag{3.9}$$

Equation (3.9) is an equation where the unknown parameters appear linearly, *i.e.*,

$$L^{-1}[y_p] = \theta^{*T}\psi \tag{3.10}$$

where θ^* is a properly defined $((n+m\rho_{\max})\times p)$ matrix of unknown parameters and ψ a $(n+m\rho_{\max})$ regressor vector defined as

$$\psi^T = \begin{bmatrix} \frac{u_1}{(s+a)} & \cdots & \frac{u_1}{(s+a)^{\rho_{\max}}} & \frac{u_2}{(s+a)} & \cdots & \frac{u_m}{(s+a)^{\rho_{\max}}} \\ \frac{y_{p1}}{(s+a)^{\rho_{\max}-\rho_1+1}} & \cdots & \frac{y_{p1}}{(s+a)^{\rho_{\max}}} & \frac{y_{p2}}{(s+a)^{\rho_{\max}-\rho_2+1}} & \cdots & \frac{y_{pp}}{(s+a)^{\rho_{\max}}} \end{bmatrix} \tag{3.11}$$

Then, the following linear equation error can be defined

$$e_2 = \theta^T\psi - L^{-1}[y_p] = (\theta^T - \theta^{*T})\psi = \phi^T\psi \tag{3.12}$$

Note that the parameters θ_{ij}^*, $i = 1, 1+\rho_{\max}, \ldots, 1+(m-1)\rho_{\max}$, j such that $\rho_j < \rho_{\max}$, are always equal to zero and do not need to be identified. The number of parameters to identify

$$p(n+m\rho_{\max}) - m(p-1) \leq N_\theta \leq p(n+m\rho_{\max})$$

Unfortunately, since the estimate of $D_L(s)$ is generally not row reduced, the conditions on the row degrees of $N_L(s)$ are not sufficient to guarantee that the estimate of the system is strictly

proper. To guarantee strict properness, the linear estimation algorithm must be modified by adding constraints on the estimates of the parameters. An example of such modifications can be found in Correa & Glover [52, 53].

Finally, it should be pointed out that other definitions for the filters $\Lambda(s)$ and $L(s)$ can be used without altering the conclusions of this work. Our filter was intentionally selected to simplify notation. Any filters of the type $\Lambda(s) = \text{diag}\{\lambda_i(s)\}$ with $\lambda_i(s)$ stable and $\partial\lambda_i \geq \max(\rho_i, \rho_{\max} - 1)$ could be used instead. The same is true for the regressor vector ψ. Other filtered signals could be selected to form the regressor vector as long as there exists a unique matrix θ^* such that equation (3.10) is verified.

Identification of canonical MFD

A similar scheme can be used for the identification of canonical left MFD than for the identification of pseudo-canonical left MFD's. A difference is that the row degrees of the matrix $D_L(s)$ are also equal to the observability indices ν_i and, consequently, less parameters need to be estimated. Also, since the estimated matrix $D_L(s)$ is row reduced, the estimated system is guaranteed to be strictly proper. So, assume that $\{N_L(s), D_L(s)\}$ is a canonical left MFD of the strictly proper transfer function matrix P(s), with observability indices $\{\nu_i\}$. Let $\Lambda(s)$ be a filter matrix defined by $\Lambda(s) = \text{diag}\{(s+a)^{\nu_i}\}$ with $a > 0$. Note that any filters of the type $\Lambda(s) = \text{diag}\{\lambda_i(s)\}$ with $\lambda_i(s)$ stable and $\partial\lambda_i \geq \nu_i$ could also be used. Then

$$y_p = \Lambda^{-1} N_L[u] - \Lambda^{-1}(D_L - \Lambda)[y_p] \tag{3.13}$$

which can be more compactly rewritten

$$y_{p_i} = \theta_i^{*T} \psi_i \qquad i = 1, \ldots, p \tag{3.14}$$

where θ_i^* and $\psi_i, i = 1, \ldots, p$ are respectively parameter vectors and regressor vectors of dimension $k_i \leq n + m\nu_i$ (with $k_{\max} = n + m\nu_{\max}$). We let ψ_i, $i = 1, \ldots, p$ be given by

$$\begin{aligned}\psi_i^T = \Big[\; & \tfrac{u_1}{(s+a)} \quad \cdots \quad \tfrac{u_1}{(s+a)^{\nu_i}} \quad \tfrac{u_2}{(s+a)} \quad \cdots \quad \tfrac{u_m}{(s+a)^{\nu_i}} \quad \tfrac{y_{p_1}}{(s+a)^{\nu_i - \min((\nu_i+1),\nu_1)+1}} \quad \cdots \quad \tfrac{y_{p_1}}{(s+a)^{\nu_i}} \\ & \tfrac{y_{p_2}}{(s+a)^{\nu_i - \min((\nu_i+1),\nu_2)+1}} \quad \cdots \quad \tfrac{y_{p_{i-1}}}{(s+a)^{\nu_i}} \quad \tfrac{y_{p_i}}{(s+a)} \quad \cdots \quad \tfrac{y_{p_i}}{(s+a)^{\nu_i}} \quad \tfrac{y_{p_{i+1}}}{(s+a)^{\nu_i - \min(\nu_i,\nu_{i+1})+1}} \quad \cdots \quad \tfrac{y_{p_p}}{(s+a)^{\nu_i}} \; \Big]\end{aligned} \tag{3.15}$$

Therefore, the following error equations can be derived

$$\begin{aligned} e_{2_i} &= \theta_i^T \psi_i - y_{p_i} = (\theta_i^T - \theta_i^{*T})\psi_i = \phi_i^T \psi_i \qquad i = 1, \ldots, p \\ \Leftrightarrow \qquad e_2 &= \text{block diag}\{\theta_i^T \psi_i\} - y_p = \theta^T \psi - y_p = \phi^T \psi \end{aligned} \tag{3.16}$$

where θ_i is the estimate of θ_i^*. The number of parameters to identify is smaller

$$p^2 \nu_{\min} + mn \leq N_\theta \leq (p+m)n$$

Also, N_θ is generally smaller (at most equal) than in the pseudo-canonical case. For example, if $p = 3$, $m = 2$, $n = 6$, $\{\nu_i\} = \{3, 1, 2\}$, and $\{\rho_i\} = \{\nu_i\}$, then $N_\theta = 28$ in the canonical case and $N_\theta = 32$ in the pseudo-canonical case. Finally, note that $N_\theta = p(p+m)\nu_{\max}$ if the observability indices are all equal and that $N_\theta = 2n$ in the SISO case, as expected.

3.2.4 Comparison

This section makes a general comparison of the identifiable parameterizations presented in this chapter in terms of necessary *a priori* information and number of parameters to identify. Table 3.1 summarizes the previous results concerning the *a priori* information and the number of adaptive parameters required by the different algorithms. Uniqueness is directly related to the

Algorithm	*A priori* information	N_θ
Direct MRAC + uniqueness	$H(s), \{\nu_i\}$	$p^2\nu_{\max} + pn$
Direct MRAC + uniqueness	$H(s), \{\rho_i\}$	$p^2\rho_{\max} + pn$
Direct PPAC + uniqueness	$\{\mu_i\}, \{\nu_i\}$	$\leq 2m^2(\nu_{\max} - 1) + m(m\mu_{\max} + n)$ $\geq 2m^2(\nu_{\max} - 1) + 2mn$
Direct PPAC + uniqueness	$\{\mu_i\}, \{\rho_i\}$	$\leq 2m^2(\rho_{\max} - 1) + m(m\mu_{\max} + n)$ $\geq 2m^2(\rho_{\max} - 1) + 2mn$
Rec. ident. + uniqueness	$\{\nu_i\}$	$\leq (p + m)n$ $\geq p^2\nu_{\min} + mn$
Rec. ident. + uniqueness	$\{\rho_i\}$	$\leq p(n + m\rho_{\max})$ $\geq p(n + m\rho_{\max}) - (p - 1)m$

Table 3.1: *A priori* information and number of parameters in identifiable parameterizations

knowledge of the observability indices (or a set of pseudo-observability indices), a fact that can be traced back to Lemma 3.1. Table 3.1, underlines the fact that uniqueness (identifiability) is obtained through increased *a priori* information, *i.e.*, the knowledge of the observability indices instead of an upper bound on the observability index. Only the *a priori* information relevant to the *parameterization* is shown in Table 3.1. Other assumptions may be needed for stability considerations. These assumptions are summarized in Table 2.2.

3.3 Frequency domain conditions

In the previous sections, we presented identifiable parameterizations for the three schemes, direct MRAC, direct PPAC, and recursive identification. Now, parameter convergence can be guaranteed if the regressor vector ψ is SE or PE (*cf.* Definitions 2.2), 2.1), and Lemma 2.3). If the regressor vector ψ is SE or PE, then $\lim_{t\to\infty} \theta = \theta^*$. With nonidentifiable parameterizations,

the regressor ψ may *never* be SE or PE. The problem with such a condition is that it constrains time signals that are interrelated and not freely available to the designer. In particular, the condition may never be satisfied, as happens when the parameterization is not identifiable. Therefore, parameter convergence conditions are not of any real use unless they are translated into conditions on the external inputs of the adaptive system. This has been done, using generalized harmonic analysis, by Boyd & Sastry [66] in the SISO case and by de Mathelin & Bodson [65] in recursive multivariable identification. Similarly, we translate in this section the PE condition on ψ into simple conditions on the reference input. Note that guaranteeing the PE condition is more interesting than guaranteeing the SE condition since the PE condition provides exponential convergence (*cf.* Lemma 2.3 for the least-squares with covariance resetting, and Anderson [68]), while the SE conditions provides only asymptotic convergence.

3.3.1 MRAC convergence results

We use the same approach than Boyd & Sastry [66] and de Mathelin & Bodson [65], to obtain the following results.

Theorem 3.1 : Frequency domain conditions for parameter convergence - Direct MRAC with K_p known

Let the inputs r_i be stationary and uncorrelated (different frequencies in each input) and assume that the observability indices $\{\nu_i\}$ are known, the corresponding identifiable parameterization is used, and the conditions of Theorem 2.1 are satisfied.

1. If, for all i, the support of the spectral measure of the i-th input, $S_{r_i}(\omega)$ contains at least $n + \nu_{\max} - p$ different values of ω, then as $t \to \infty$, $\theta \to \theta^*$, $\forall$ initial conditions.

2. If the support of the spectral measure of the inputs taken altogether, $S_R(\omega)$ does not contain at least $n + p(\nu_{\max} - 1)$ different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

Corollary 3.1

Let the inputs r_i be stationary and uncorrelated (different frequencies in each input) and assume that a set of pseudo-observability indices $\{\rho_i\}$ is known, the corresponding identifiable parameterization is used, and the conditions of Theorem 2.1 are satisfied.

1. If, for all i, the support of the spectral measure of the i-th input, $S_{r_i}(\omega)$ contains at least $n + \rho_{\max} - p$ different values of ω, then as $t \to \infty$, $\theta \to \theta^*$, $\forall$ initial conditions.

2. If the support of the spectral measure of the inputs taken altogether, $S_R(\omega)$ does not contain at least $n + p(\rho_{\max} - 1)$ different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

Theorem 3.2 : Frequency domain conditions for parameter convergence - Direct MRAC with K_p unknown

Let the inputs r_i be stationary and uncorrelated (different frequencies in each input) and assume that the observability indices $\{\nu_i\}$ are known, the corresponding identifiable parameterization is used, and the conditions of Theorem 2.2 or Theorem 4.1 are satisfied.

1. If, for all i, the support of the spectral measure of the i-th input, $S_{r_i}(\omega)$ contains at least $n + \nu_{\max} - p + 1$ different values of ω, then as $t \to \infty$, $\theta \to \theta^*$, $\forall$ initial conditions.

2. If the support of the spectral measure of the inputs taken altogether, $S_R(\omega)$ does not contain at least $n + p\nu_{\max}$ different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

Corollary 3.2

Let the inputs r_i be stationary and uncorrelated (different frequencies in each input) and assume that a set of pseudo-observability indices $\{\rho_i\}$ is known, the corresponding identifiable parameterization is used, and the conditions of Theorem 2.2 or Theorem 4.1 are satisfied.

1. If, for all i, the support of the spectral measure of the i-th input, $S_{r_i}(\omega)$ contains at least $n + \rho_{\max} - p + 1$ different values of ω, then as $t \to \infty$, $\theta \to \theta^*$, $\forall$ initial conditions.

2. If the support of the spectral measure of the inputs taken altogether, $S_R(\omega)$ does not contain at least $n + p\rho_{\max}$ different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

The proof of Theorem 3.1 is a trivial simplification of the proof of Theorem 3.2 given below. The proof of Corollaries 3.1 and 3.2 are easily derived from the proof of Theorems 3.1 and 3.2 by simply using the identifiable controller parameterization based on the pseudo-observability indices instead of the parameterization based on the observability indices. For the interpretation of the results, it should be noted that one sinusoid, $\sin(\omega_0 t)$, contributes to two frequency components, at $+\omega_0$ and $-\omega_0$.

Proof of Theorem 3.2

Since $e_\psi \in L_2$ and $\lim_{t\to\infty} e_\psi = 0$ (*cf.* Theorems 2.1, 2.2, and 4.1)

$$\psi \ \text{PE} \quad \Leftrightarrow \quad \psi_m \ \text{PE}$$

(*cf.* Lemma 2.6.6 in Sastry & Bodson [3]). Since the reference input r is assumed stationary

$$\psi_m \ \text{PE} \quad \Leftrightarrow \quad R_{\psi_m}(0) > 0$$

where $R_{\psi_m}(t)$ is the autocorrelation of $\psi_m(t)$. By the properties of the autocorrelation

$$R_{\psi_m}(0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} H_{\psi_m r}(j\omega) S_R(\omega) H^{*T}_{\psi_m r}(j\omega)\, d\omega$$

where $S_R(\omega)$ is the spectral measure of r, *e.g.*, $S_R(\omega) = \mathrm{diag}\{S_{r_i}(\omega)\}$ if the $m = p$ reference inputs r_i are uncorrelated (different frequencies by input signal). Since the identifiable parameterization is used (see Section 3.2.1), the $(n + p\nu_{\max}) \times p$ transfer function matrix $H_{\psi_m r}$ has the following form (*cf.* equations (2.13) and (2.30))

$$H_{\psi_m r}(s) = \begin{bmatrix} \mathrm{diag}\{\frac{1}{l(s)}\} \\ \frac{1}{l(s)\lambda(s)} P^{-1}(s) H(s) \\ \vdots \\ \frac{s^{\nu_{\max}-2}}{l(s)\lambda(s)} P^{-1}(s) H(s) \\ \frac{1}{l(s)\lambda(s)} T^1 H(s) \\ \vdots \\ \frac{s^{\nu_{\max}-2}}{l(s)\lambda(s)} T^{\nu_{\max}-1} H(s) \\ \frac{1}{l(s)} T^{\nu_{\max}} H(s) \end{bmatrix}$$

where T^i is equal to the identity matrix I without the rows I_j such that $\nu_j < i$. For example, if $p = 3, n = 6, \{\nu_i\} = \{3, 1, 2\}$ then

$$T^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad T^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad T^3 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

Part 1: To establish the first part of the theorem we use the following intermediate result: Let $H_{\psi_m r_i}$ denote the i-th column of $H_{\psi_m r}$, and let $N_f = n + \nu_{\max} - p + 1$. Then, for all set of *distinct* $\{s_{ij} \in \bar{C}\}$, $i = 1, \ldots, p$, $j = 1, \ldots, N_f$, such that $H_{\psi_m r}(s_{ij})$ is well defined, the rows of the matrix

$$\begin{bmatrix} H_{\psi_m r_1}(s_{11}) & \ldots & H_{\psi_m r_1}(s_{1N_f}) & H_{\psi_m r_2}(s_{21}) & \ldots & H_{\psi_m r_p}(s_{pN_f}) \end{bmatrix}$$

are linearly independent in $C^{pN_f = N_\theta}$.

This intermediate result is proved in the following manner. Suppose that the proposition was not true. Then, there would exist a vector $k \in C^{n+p\nu_{\max}} \neq 0$, such that:

$$k^T \begin{bmatrix} H_{\psi_m r_1}(s_{11}) & \ldots & H_{\psi_m r_p}(s_{pN_f}) \end{bmatrix} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix}$$

Let $k^T = \begin{bmatrix} k_1^T & \cdots & k_{2\nu_{\max}}^T \end{bmatrix}$, where k_i is an $(p = m)$ dimensional vector for $i = 1, \ldots, \nu_{\max}$ and an $(\leq p)$ dimensional vector for $i = \nu_{\max} + 1, \ldots, 2\nu_{\max}$ (*i.e.*, with a dimension equal to the

number of rows of $T^{i-\nu_{\max}}$). Then, given the definition of $H_{\psi_m r}$

$$
\begin{aligned}
0 \;=\;& k_{1_i}\frac{1}{l(s_{ij})} + k_2^T \frac{1}{l(s_{ij})\lambda(s_{ij})}(P^{-1}H)_i(s_{ij}) + \ldots + k_{\nu_{\max}}^T \frac{s_{ij}^{\nu_{\max}-2}}{l(s_{ij})\lambda(s_{ij})}(P^{-1}H)_i(s_{ij}) + \\
& k_{\nu_{\max}+1}^T \frac{1}{l(s_{ij})\lambda(s_{ij})} T^1 H_i(s_{ij}) + \ldots + k_{2\nu_{\max}-1}^T \frac{s_{ij}^{\nu_{\max}-2}}{l(s_{ij})\lambda(s_{ij})} T^{\nu_{\max}-1} H_i(s_{ij}) \\
& + k_{2\nu_{\max}}^T \frac{1}{l(s_{ij})} T^{\nu_{\max}} H_i(s_{ij}) \qquad \forall\, i = 1,\ldots,p \text{ and } j = 1,\ldots,N_f \qquad (3.17)
\end{aligned}
$$

where $(P^{-1}H)_i(s)$ is the i-th column of $(P^{-1}H)(s)$, $H_i(s)$ the i-th column of $H(s)$, and k_{1_i} the i-th element of the vector k_1.

Given the definition of the Hermite normal form $H(s)$ (Definition 1.4), $\exists h_i, 1 \le h_i \le r_{\max}$ such that the greatest common denominator of the elements of $H_i(s)$ is equal to $(s+a)^{h_i}$. Since $\lim_{s\to\infty} P^{-1}(s)H(s) = K_p^{-1}$ nonsingular, if $\{N_L, D_L\}$ are coprime then

$$\partial \det N_L = \partial \det D_L - \partial \det H^{-1}(s) = n - \sum_{k=1}^{p} r_k$$

Therefore, if we define $q_i(s)$ as the smallest common denominator of $(P^{-1}H)_i(s) = N_L^{-1}(s)D_L(s)$ $H_i(s)$, then $q_i(s) = \det N_L(s)(s+a)^{h_i}$ and

$$\partial q_i(s) = n + h_i - \sum_{k=1}^{p} r_k \le n - p + 1$$

since $r_k \ge 1\ \forall k$. Now, if we multiply both sides of (3.17) by $l(s_{ij})\lambda(s_{ij})q_i(s_{ij})$

$$
\begin{aligned}
0 \;=\;& k_{1_i}\lambda(s_{ij})q_i(s_{ij}) + k_2^T(P^{-1}H)_i(s_{ij})q_i(s_{ij}) + \ldots + k_{\nu_{\max}}^T s_{ij}^{\nu_{\max}-2}(P^{-1}H)_i(s_{ij})q_i(s_{ij}) \\
& + k_{\nu_{\max}+1}^T T^1 H_i(s_{ij})q_i(s_{ij}) + \ldots + k_{2\nu_{\max}-1}^T s_{ij}^{\nu_{\max}-2} T^{\nu_{\max}-1} H_i(s_{ij})q_i(s_{ij}) \\
& + k_{2\nu_{\max}}^T \lambda(s_{ij}) T^{\nu_{\max}} H_i(s_{ij})q_i(s_{ij}) \qquad \forall\, i = 1,\ldots,p \text{ and } j = 1,\ldots,N_f
\end{aligned}
$$

These are p polynomial equations, each with degree $< \nu_{\max} + n - p + 1$. Consequently, if $N_f \ge \nu_{\max} + n - p + 1$, these equations cannot be verified unless they are verified $\forall\ s$ or equivalently if

$$
\begin{aligned}
0 \;=\;& k_1^T \frac{1}{l(s)} + k_2^T \frac{1}{l(s)\lambda(s)} P^{-1}(s)H(s) + \ldots + k_{\nu_{\max}}^T \frac{s^{\nu_{\max}-2}}{l(s)\lambda(s)} P^{-1}(s)H(s) + \\
& k_{\nu_{\max}+1}^T \frac{1}{l(s)\lambda(s)} T^1 H(s) + \ldots + k_{2\nu_{\max}-1}^T \frac{s^{\nu_{\max}-2}}{l(s)\lambda(s)} T^{\nu_{\max}-1} H(s) + k_{2\nu_{\max}}^T \frac{1}{l(s)} T^{\nu_{\max}} H(s)
\end{aligned}
$$

or

$$
\begin{aligned}
0 \;=\;& k_1^T H^{-1}(s)P(s) + \frac{1}{\lambda(s)}[k_2^T + s k_3^T + \ldots + s^{\nu_{\max}-2} k_{\nu_{\max}}^T] \\
& + \frac{1}{\lambda(s)}[k_{\nu_{\max}+1}^T T^1 + \ldots + s^{\nu_{\max}-2} k_{2\nu_{\max}-1}^T T^{\nu_{\max}-1} + \lambda(s) k_{2\nu_{\max}}^T T^{\nu_{\max}}] P(s)
\end{aligned}
$$

If we take the $\lim_{s\to\infty}$, we obtain

$$k_1^T K_p = 0 \quad \Leftrightarrow \quad k_1 = 0$$

so that

$$\begin{aligned} 0 \;=\; & [k_2^T + s k_3^T + \ldots + s^{\nu_{\max}-2} k_{\nu_{\max}}^T] D_R(s) \\ & + [k_{\nu_{\max}+1}^T T^1 + \ldots + s^{\nu_{\max}-2} k_{2\nu_{\max}-1}^T T^{\nu_{\max}-1} + \lambda(s) k_{2\nu_{\max}}^T T^{\nu_{\max}}] N_R(s) \end{aligned}$$

and by the null matrix condition (Lemma 3.1)

$$k_i = 0 \quad \forall\, i = 2, \ldots, 2\nu_{\max} \qquad \Rightarrow \qquad k = 0$$

which proves our intermediate result.

Now, suppose $R_{\psi_m}(0) \not> 0$. Then, there exists a vector $k \in \Re^{n+p\nu_{\max}}, k \neq 0$ such that $\int_{-\infty}^{+\infty} k^T H_{\psi_m r}(j\omega) S_R(\omega) H_{\psi_m r}^{*T}(j\omega) k \, d\omega = 0$. From our intermediate result, $k^T H_{\psi_m r_i}(j\omega_i) = 0$ for $i = 1, \ldots, p$ is not possible for $n + \nu_{\max} - p + 1$ different values of ω_i. Therefore $S_{r_i}(\omega) > 0$ for less than $n + \nu_{\max} - p + 1$ values of ω, which contradicts the hypothesis of the first part of the theorem.

Part 2: Let $N_f < n + p\nu_{\max}$. Then, for all set of $\{s_i \in C\}$, $i = 1, \ldots, N_f$, such that $H_{\psi_m r}(s_i)$ is well defined, the rows of the matrix

$$\left[\; H_{\psi_m r_1}(s_1) \quad \ldots \quad H_{\psi_m r_p}(s_{N_f}) \;\right]$$

are linearly dependent in C^{N_f}, since the number of rows $(n + p\nu_{\max}) > N_f$ the number of columns. Consequently, if there are less than $n + p\nu_{\max}$ different values of ω in the support of $S_R(\omega)$, then there exists a vector $k \in \Re^{n+p\nu_{\max}}, k \neq 0$ such that $\int_{-\infty}^{+\infty} k^T H_{\psi_m r}(j\omega)\; S_R(\omega)\; H_{\psi_m r}^{*T}(j\omega) k \, d\omega = 0$. Therefore, $R_{\psi_m}(0) \not> 0$. This implies that $\phi(t)$ does not converge to zero, unless the initial error $\phi(0)$ lies exactly in the null space of $R_{\psi_m}(0)$ (*cf.* Sastry & Bodson [3]). □

3.3.2 PPAC convergence results

Similarly to the MRAC case, PE conditions can be translated into conditions on the reference input.

Theorem 3.3 : Frequency domain conditions for parameter convergence - Direct PPAC

Let the inputs r_i be stationary and uncorrelated (different frequencies in each input) and assume that the observability indices $\{\nu_i\}$ are known, the corresponding identifiable parameterization is used, and conditions of Theorem 2.3 are satisfied.

1. If, for all i, the support of the spectral measure of the i-th input, $S_{r_i}(\omega)$ contains at least $(m+1)(\nu_{\max}-1)+m\mu_{\max}+\mu_i$ different values of ω, then as $t \to \infty$, $\theta \to \theta^*$ $\forall$ initial conditions.

2. If the support of the spectral measure of the inputs taken altogether, $S_R(\omega)$ does not contain at least $n+2m(\nu_{\max}-1)+m\mu_{\max}$ different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

Corollary 3.3

Let the inputs r_i be stationary and uncorrelated (different frequencies in each input) and assume that a set of pseudo-observability indices $\{\rho_i\}$ is known, the corresponding identifiable parameterization is used, and conditions of Theorem 2.3 are satisfied.

1. If, for all i, the support of the spectral measure of the i-th input, $S_{r_i}(\omega)$ contains at least $(m+1)(\rho_{\max}-1)+m\mu_{\max}+\mu_i$ different values of ω, then as $t \to \infty$, $\theta \to \theta^*$ $\forall$ initial conditions.

2. If the support of the spectral measure of the inputs taken altogether, $S_R(\omega)$ does not contain at least $n+2m(\rho_{\max}-1)+m\mu_{\max}$ different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

The proof of Corollary 3.3 is easily derived from the proof of Theorem 3.3 by simply using the identifiable controller parameterization based on the pseudo-observability indices instead of the parameterization based on the observability indices.

Proof of Theorem 3.3

Part 1: We prove the result by contradiction. Suppose that ψ is not PE. Since the conditions of Theorem 2.3 are satisfied, we have that $\lim_{t\to\infty}\theta(t)=\theta_\infty$ and from the proof of Theorem 2.3,

$$\psi-\psi_\infty = H_{\psi_\infty r}[(\tilde{\theta}-\tilde{\theta}_\infty)^T w]$$

Therefore, since $H_{\psi_\infty r}$ is strictly proper and stable, $\lim_{t\to\infty}(\psi-\psi_\infty)=0$ and

$$\psi \ \text{ not PE} \quad \Rightarrow \quad \psi_\infty \ \text{ not PE}$$

Since the reference input r is assumed stationary

$$\psi_\infty \ \text{ not PE} \quad \Leftrightarrow \quad R_{\psi_\infty}(0) \not> 0$$

where $R_{\psi_\infty}(t)$ is the autocorrelation of $\psi_\infty(t)$ and

$$R_{\psi_\infty}(0)=\frac{1}{2\pi}\int_{\infty}^{\infty} H_{\psi_\infty r}(j\omega)S_R(\omega)H^{*T}_{\psi_\infty r}(j\omega)\,d\omega$$

where $S_R(\omega)$ is the spectral measure of r. See the identifiable parameterization is used (see Section 3.2.2), the $(2m(\nu_{\max}-1)+m\mu_{\max}+n)\times m$ transfer function matrix $H_{\psi_\infty r}$ has the following form (*cf.* equations (2.23) and (7.1))

$$H_{\psi_\infty r}(s) = \begin{bmatrix} \frac{1}{l(s)\gamma(s)} S^1 D_R(s) Q^{-1}(s)\Lambda(s) \\ \vdots \\ \frac{s^{\mu_{\max}-\mu_{\min}+\nu_{\max}-2}}{l(s)\gamma(s)} S^{\mu_{\max}-\mu_{\min}+\nu_{\max}-1} D_R(s) Q^{-1}(s)\Lambda(s) \\ \frac{1}{l(s)\gamma(s)} T^1 N_R(s) Q^{-1}(s)\Lambda(s) \\ \vdots \\ \frac{s^{\nu_{\max}-1}}{l(s)\gamma(s)} T^{\nu_{\max}} N_R(s) Q^{-1}(s)\Lambda(s) \\ \frac{1}{l(s)} D_R(s) Q^{-1}(s)\Lambda(s) \\ \vdots \\ \frac{s^{\nu_{\max}-2}}{l(s)} D_R(s) Q^{-1}(s)\Lambda(s) \\ \frac{1}{l(s)} T^1 N_R(s) Q^{-1}(s)\Lambda(s) \\ \vdots \\ \frac{s^{\nu_{\max}-1}}{l(s)} T^{\nu_{\max}} N_R(s) Q^{-1}(s)\Lambda(s) \end{bmatrix}$$

where $Q(s) = (\Lambda(s) - \bar{C}_\infty(s))D_R(s) - \bar{D}_\infty(s)N_R(s)$, T^i is equal to the $(p\times p)$ identity matrix I_p without the rows I_{p_j} such that $\nu_j < i$, and S^i is equal to the $(m\times m)$ identity matrix I_m without the rows I_{m_j} such that $\mu_{\max}-\mu_j+\nu_{\max}-1 < i$. For example, if $m=3, n=6, \nu_{\max}=2, \{\mu_i\}=\{3,1,2\}$ then

$$S^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad S^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad S^3 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

Note that $Q^{-1}(s)\Lambda(s) = D_M^{-1}(s)$ if $\theta_\infty = \theta^*$ and $\psi_\infty = \psi_m$.

We need now to establish the following intermediate result:
Let $H_{\psi_\infty r_i}$ denote the i-th column of $H_{\psi_\infty r}$, and let $N_{f_i} = (m+1)(\nu_{\max}-1)+m\mu_{\max}+\mu_i$. Then, for all set of *distinct* $\{s_{ij}\in C\}$, $i=1,\ldots,m$, $j=1,\ldots,N_{f_i}$, such that $H_{\psi_\infty r}(s_{ij})$ is well defined, the rows of the matrix

$$\begin{bmatrix} H_{\psi_\infty r_1}(s_{11}) & \ldots & H_{\psi_\infty r_1}(s_{1N_{f_1}}) & H_{\psi_\infty r_2}(s_{21}) & \ldots & H_{\psi_\infty r_m}(s_{mN_{f_m}}) \end{bmatrix}$$

are linearly independent in $C^{\sum_{i=1}^m N_{f_i}}$.

This intermediate result is proved in the same manner as for Theorem 3.2. Suppose that the proposition was not true. Then, there would exist a vector $k \in C^{2m(\nu_{\max}-1)+m\mu_{\max}+n} \neq 0$, such that:

$$k^T \left[\begin{array}{ccc} H_{\psi_\infty r_1}(s_{11}) & \dots & H_{\psi_\infty r_m}(s_{mN_{fm}}) \end{array} \right] = \left[\begin{array}{ccc} 0 & \dots & 0 \end{array} \right]$$

Let $k^T = \left[\begin{array}{ccc} k_1^T & \cdots & k^T_{\mu_{\max}-\mu_{\min}+4\nu_{\max}-2} \end{array} \right]$, where the subvectors k_i are properly dimensioned so that the previous equation can be rewritten as

$$\begin{aligned}
0 \;=\; & k_1^T \frac{1}{l(s_{ij})d_i(s_{ij})} S^1 D_R(s_{ij}) Q_i^{-1}(s_{ij}) + \dots \\
& + k^T_{\mu_{\max}-\mu_{\min}+\nu_{\max}-1} \frac{s_{ij}^{\mu_{\max}-\mu_{\min}+\nu_{\max}-2}}{l(s_{ij})d_i(s_{ij})} S^{\mu_{\max}-\mu_{\min}+\nu_{\max}-1} D_R(s_{ij}) Q_i^{-1}(s_{ij}) \\
& + k^T_{\mu_{\max}-\mu_{\min}+\nu_{\max}} \frac{1}{l(s_{ij})d_i(s_{ij})} T^1 N_R(s_{ij}) Q_i^{-1}(s_{ij}) + \dots \\
& + k^T_{\mu_{\max}-\mu_{\min}+2\nu_{\max}-1} \frac{s_{ij}^{\nu_{\max}-1}}{l(s_{ij})d_i(s_{ij})} T^{\nu_{\max}} N_R(s_{ij}) Q_i^{-1}(s_{ij}) \\
& + k^T_{\mu_{\max}-\mu_{\min}+2\nu_{\max}} \frac{\lambda_i(s_{ij})}{l(s_{ij})} D_R(s_{ij}) Q_i^{-1}(s_{ij}) + \dots \\
& + k^T_{\mu_{\max}-\mu_{\min}+3\nu_{\max}-2} \frac{s_{ij}^{\nu_{\max}-2}\lambda_i(s_{ij})}{l(s_{ij})} D_R(s_{ij}) Q_i^{-1}(s_{ij}) \\
& + k^T_{\mu_{\max}-\mu_{\min}+3\nu_{\max}-1} \frac{\lambda_i(s_{ij})}{l(s_{ij})} T^1 N_R(s_{ij}) Q_i^{-1}(s_{ij}) + \dots \\
& + k^T_{\mu_{\max}-\mu_{\min}+4\nu_{\max}-2} \frac{s_{ij}^{\nu_{\max}-1}\lambda_i(s_{ij})}{l(s_{ij})} T^{\nu_{\max}} N_R(s_{ij}) Q_i^{-1}(s_{ij}) \\
& \qquad \forall\, i = 1, \dots, m \;\text{ and }\; j = 1, \dots, N_{f_i}
\end{aligned} \tag{3.18}$$

where $Q_i^{-1}(s)$ is the i-th column of $Q^{-1}(s)$. Suppose that $q_i(s)$ is the smallest common denominator of $Q_i^{-1}(s)$, then $\partial q_i \leq \partial \det(Q) \leq m(\mu_{\max} + \nu_{\max} - 1)$ and $\partial(D_R Q_i^{-1} q_i \lambda_i) \leq \partial q_i \leq m(\mu_{\max} + \nu_{\max} - 1)$. Note that $q_i(s) = \gamma(s)$ if $\psi_\infty = \psi_m$.

Now, if we multiply both sides of (3.18) by $l(s_{ij})d_i(s_{ij})q_i(s_{ij})$, we obtain a system of m polynomial equations with degree $< (m+1)(\nu_{\max} - 1) + m\mu_{\max} + \mu_i$. Consequently, if $N_{f_i} \geq (m+1)(\nu_{\max} - 1) + m\mu_{\max} + \mu_i$, $\forall\, i = 1, \dots, m$, these equations cannot be verified unless they are verified $\forall\, s$ or equivalently if

$$\begin{aligned}
& \frac{1}{\gamma(s)} \Big([k_1^T S^1 + \dots + k^T_{\mu_{\max}-\mu_{\min}+\nu_{\max}-1} s^{\mu_{\max}-\mu_{\min}+\nu_{\max}-2} S^{\mu_{\max}-\mu_{\min}+\nu_{\max}-1}] D_R(s) \\
& + [k^T_{\mu_{\max}-\mu_{\min}+\nu_{\max}} T^1 + \dots + k^T_{\mu_{\max}-\mu_{\min}+2\nu_{\max}-1} s^{\nu_{\max}-1} T^{\nu_{\max}}] N_R(s) \Big) \\
& = -\Big([k^T_{\mu_{\max}-\mu_{\min}+2\nu_{\max}} + \dots + k^T_{\mu_{\max}-\mu_{\min}+3\nu_{\max}-2} s^{\nu_{\max}-2}] D_R(s)
\end{aligned}$$

$$+ [k^T_{\mu_{\max}-\mu_{\min}+3\nu_{\max}-1} T^1 + \ldots + k^T_{\mu_{\max}-\mu_{\min}+4\nu_{\max}-2} s^{\nu_{\max}-1} T^{\nu_{\max}}] N_R(s)\Big)$$

The left side of this equation is strictly proper and the right side is polynomial, so that both sides must be identically zero. Then, by the null matrix condition (Lemma 3.1)

$$k_i = 0 \quad \forall\, i = 1, \ldots, \mu_{\max} - \mu_{\min} + 4\nu_{\max} - 2 \qquad \Rightarrow \qquad k = 0$$

which proves the intermediate result.

Since $R_{\psi_\infty}(0) \not> 0$, there exists a vector $k \in \Re^{2m(\nu_{\max}-1)+m\mu_{\max}+n}$, $k \neq 0$ such that $\int_{-\infty}^{+\infty} k^T H_{\psi_\infty r}(j\omega)\ S_R(\omega)\ H^{*T}_{\psi_\infty r}(j\omega) k\ d\omega = 0$. Since, we proved in the intermediate result that $k^T H_{\psi_\infty r_i}(j\omega_i) = 0$ for $i = 1, \ldots, m$ is not possible for $(m+1)(\nu_{\max} - 1) + m\mu_{\max} + \mu_i$ different values of ω_i, then $S_{r_i}(\omega) > 0$ for less than $(m+1)(\nu_{\max} - 1) + m\mu_{\max} + \mu_i$ values of ω. This contradicts the hypothesis of the first part of the theorem.

Part 2: We also prove this result by contradiction. If ψ was PE, then $\theta_\infty = \theta^*$ and $\psi_\infty = \psi_m$. Since $\lim_{t\to\infty} (\psi - \psi_\infty) = 0$, ψ_m would also be PE. However, let $N_f < 2m(\nu_{\max}-1) + m\mu_{\max} + n$. Then, for all set of $\{s_i \in C\},\ i = 1, \ldots, N_f$, such that $H_{\psi_m r}(s_i)$ is well defined, the rows of the matrix

$$\left[\ H_{\psi_m r_1}(s_1) \quad \ldots \quad H_{\psi_m r_m}(s_{N_f})\ \right]$$

are linearly dependent in C^{N_f} since the number of rows $(2m(\nu_{\max} - 1)m\mu_{\max} + n) > N_f$ the number of columns. Consequently, if there are less than $2m(\nu_{\max} - 1) + m\mu_{\max} + n$ different values of ω in the support of $S_R(\omega)$, then there exists a vector $k \in \Re^{2m(\nu_{\max})+m\mu_{\max}+n}$, $k \neq 0$ such that $\int_{-\infty}^{+\infty} k^T\ H_{\psi_m r}(j\omega)\ S_R(\omega)\ H^{*T}_{\psi_m r}(j\omega)\ k\ d\omega = 0$, and $R_{\psi_m}(0) \not> 0$. This implies that ψ_m is not PE and therefore we have reached a contradiction. In consequence, $\phi(t)$ does not converge to zero, unless the initial error $\phi(0)$ lies exactly in the null space of $R_{\psi_m}(0)$. □

3.3.3 Recursive identification convergence results

The PE condition was transformed into a condition on the reference input u in recursive multivariable identification by de Mathelin & Bodson [65]. For the sake of comparison, these results are repeated here with their extension to the identification of pseudo-canonical MFD's.

Theorem 3.4 : Frequency domain conditions for parameter convergence - Identification of canonical left MFD

Let the inputs u_i be stationary and uncorrelated (different frequencies in each input)

1. If, for all j, the support of the spectral measure of the j-th input, $S_{u_j}(\omega)$ contains at least $n + \nu_i$ different values of ω then as $t \to \infty$, $\theta_i \to \theta_i^*$, and if it contains at least $n + \nu_{\max}$ different values of ω then as $t \to \infty$, $\theta \to \theta^*$.

2. If the support of the spectral measure of the inputs taken altogether, $S_U(\omega)$ does not contain at least k_i different values of ω, or if the support of the spectral measure of one of the inputs does not contain at least ν_i different values of ω, then as $t \to \infty$, $\theta_i \not\to \theta_i^*$ for almost all initial conditions. If it does not contain at least $n + m\nu_{\max}$ different values of ω, or if the support of the spectral measure of one of the inputs does not contain at least ν_{max} different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

Theorem 3.5 : Frequency domain conditions for parameter convergence - Identification of pseudo-canonical left MFD

Let the inputs u_i be stationary and uncorrelated (different frequencies in each input)

1. If, for all i, the support of the spectral measure of the i-th input, $S_{u_i}(\omega)$ contains at least $n + \rho_{\max}$ different values of ω then as $t \to \infty$, $\theta \to \theta^*$.

2. If the support of the spectral measure of the inputs taken altogether, $S_U(\omega)$ does not contain at least $n + m\rho_{\max}$ different values of ω, or if the support of the spectral measure of one of the inputs does not contain at least ω_{max} different values of ω, then as $t \to \infty$, $\theta \not\to \theta^*$ for almost all initial conditions.

Remarks:

The proof of Theorem 3.4 can be found in in de Mathelin & Bodson [65]. The proof of Theorem 3.5 can be easily derived from the proof of Theorem 3.4. In the SISO case, the sufficient condition for parameter convergence is equal to the necessary condition. As expected, $2n$ frequency components are necessary. Consequently, the number of frequency components necessary for parameter convergence is equal to the number of parameters N_θ. This is not true for multivariable systems where less frequencies are required than the number of parameters.

Generally, the sufficient condition is different from the necessary condition (it is equal when $m = 1$). When the necessary condition is satisfied but the sufficient condition is not, parameter convergence may or may not occur depending on the value of the frequency components in the input. This peculiar phenomenon was reported in de Mathelin & Bodson [65].

It is also interesting to note that the condition for the convergence of $\theta \to \theta^*$ is identical in both cases if $\nu_{\max} = \rho_{\max}$. However, in the canonical case (with usually less parameters to identify), the convergence of the parameters of rows with smaller observability indices can be guaranteed by smaller frequency domain conditions.

3.3.4 Comparison

In the SISO case, the sufficient condition for parameter convergence is equal to the necessary condition. As expected (*cf.* Boyd & Sastry [66] and Elliot *et al.* [62]), $2n$ frequency components are necessary in the MRAC case (with unknown k_p), and $4n - 2$ in the PPAC case. Consequently, the number of frequency components necessary for parameter convergence is equal to

the number of parameters N_θ. This is not true for multivariable systems where less frequencies are required than the number of parameters. Generally, in the multivariable adaptive control case, the sufficient condition is different from the necessary condition. A similar result was found in the identification case by de Mathelin & Bodson [65]. Furthermore,

$$\text{n.c. Ident. (canonical)} = \text{n.c. direct MRAC} \leq \text{s.c. direct MRAC} \leq \text{s.c. Ident. (canonical)}$$

where Ident. stands for recursive identification, n.c. for necessary condition, and s.c. for sufficient condition summed up on all the inputs. A particularly interesting fact is that when the necessary condition is satisfied but the sufficient condition is not, parameter convergence occurs in certain cases and not in others, depending on the location of the frequency components in the input. This peculiar phenomenon proper to multivariable systems is exposed for MRAC systems through examples in the section to follow. Furthermore, it will also be shown through examples that for some systems and for some particular input signals, parameter convergence occurs in the identification case but not in the MRAC case, and *vice versa*. It is also interesting to note that there is no relationship between the particular frequencies and the zeros of the system. Indeed, there exists an infinite number of different input signals for which this particular phenomenon occurs and it occurs in the MRAC case for different signals than in the identification case.

In the direct PPAC case, the sufficient condition is equal to the necessary condition when there is only one input ($m = 1$). Also,

$$\text{n.c. Identification (canonical)} = \text{n.c. direct MRAC} \leq \text{n.c. direct PPAC}$$

and

$$\text{s.c. direct MRAC} \leq \text{s.c. Identification (canonical)} \leq \text{s.c. direct PPAC}$$

However, the necessary condition in the direct PPAC case may or may not be smaller than the sufficient condition (summed on all the inputs) in the direct MRAC case and identification case, depending on the value of m, n, $\nu_{\max}$ (or $\rho_{\max}$), and $\mu_{\max}$. Consequently, parameter convergence could occur in the direct PPAC case and not in the direct MRAC case or in the identification case and vice versa, for some systems and some particular frequencies. Generally, however, more frequencies will be needed for parameter convergence in the direct PPAC case. Intuitively, this is to be expected since the direct PPAC usually requires more parameters than the other algorithms (twice as many in the SISO case). Finally, Table 3.2 summarizes the frequency domain conditions for parameter convergence for the three different schemes under study.

3.4 Examples

3.4.1 Robustness

It is reasonable to ask why parameter convergence and identifiable parameterizations would be useful in adaptive control. The first reason is that when nonidentifiable parameterizations are

Algorithm	Sufficient condition (per input)	Necessary condition (inputs taken altogether)
Direct MRAC (K_p known)	$n + \nu_{\max} - p$	$n + p(\nu_{\max} - 1)$
Direct MRAC (K_p known)	$n + \rho_{\max} - p$	$n + p(\rho_{\max} - 1)$
Direct MRAC (K_p unknown)	$n + \nu_{\max} - p + 1$	$n + p\nu_{\max}$
Direct MRAC (K_p unknown)	$n + \rho_{\max} - p + 1$	$n + p\rho_{\max}$
Direct PPAC	$(m+1)(\nu_{\max} - 1) + m\mu_{\max} + \mu_i$	$2m(\nu_{\max} - 1) + m\mu_{\max} + n$
Direct PPAC	$(m+1)(\rho_{\max} - 1) + m\mu_{\max} + \mu_i$	$2m(\rho_{\max} - 1) + m\mu_{\max} + n$
Recursive identification	$n + \nu_{\max}$	$n + m\nu_{\max}$
Recursive identification	$n + \rho_{\max}$	$n + m\rho_{\max}$

Table 3.2: Conditions for parameter convergence

used in adaptive control, conditions of persistency of excitation on the regressor vectors cannot be satisfied in general, even with rich signals, so that the parameters will not converge to unique values. Since the set to which the parameters converge is usually unbounded, the convergence of the parameters can be very sensitive to disturbances. Small measurement noise and unmodeled dynamics can easily force convergence of the parameters to regions of the parameter space where the system becomes unstable. Conversely, with identifiable parameterizations and sufficiently rich inputs, a certain degree of robustness to noise and unmodeled dynamics is always guaranteed and the parameters will remain in the neighborhood of their nominal values (*cf.* Sastry & Bodson [3]).

Another advantage of identifiable parameterizations is that they usually require a smaller number of parameters. This number already tends to be large for MIMO systems and computational requirements grow fast with the number of parameters, especially for least-squares algorithms. Given the real-time constraints of adaptive control, the number of parameters to adapt is therefore a crucial consideration in practice.

The following simulation illustrates the fact that the schemes using identifiable parameterizations generally have better robustness properties than those using nonidentifiable parameterizations. This example is one of several simulations that we ran, adding various types of noise and unmodeled dynamics to the ideal systems. In our experiments, we found several cases where both schemes (using identifiable and nonidentifiable parameterizations) were stable, and several cases (as the one presented here) where only the scheme based on an identifiable parameterization remained stable. The reverse was not observed.

Let the nominal plant be a 2-input 2-output system of order 3, with the following transfer

function matrix

$$P(s) = \begin{bmatrix} \frac{2s^2+3s-4}{s^3+3s^2+4s+2} & \frac{s^2-s-8}{s^3+3s^2+4s+2} \\ \frac{-s^2-4s+7}{s^3+3s^2+4s+2} & \frac{s+13}{s^3+3s^2+4s+2} \end{bmatrix}$$

The observability indices are $\nu_1 = 2$ and $\nu_2 = 1$ and the canonical left MFD is

$$N_L(s) = \begin{bmatrix} 2s-1 & s-3 \\ 1 & 1 \end{bmatrix} \qquad D_L(s) = \begin{bmatrix} s^2+s+4 & 2 \\ s+3 & s+2 \end{bmatrix}$$

Since

$$\begin{aligned} \det N_L(s) &= s+2 \\ \det D_L(s) &= s^3+3s^2+4s+2 = (s+1)(s^2+2s+2) \end{aligned}$$

the system is stable and minimum phase. Assume that the MRAC algorithm is implemented with

$$\Lambda(s) = \begin{bmatrix} (s+\lambda) & 0 \\ 0 & (s+\lambda) \end{bmatrix} \quad \lambda > 0 \qquad L(s) = \begin{bmatrix} (\frac{s}{l}+1) & 0 \\ 0 & (\frac{s}{l}+1) \end{bmatrix} \quad l > 0$$

$$M(s) = H(s) = \begin{bmatrix} \frac{1}{(s+1)} & 0 \\ 0 & \frac{1}{(s+1)} \end{bmatrix}$$

Identifiable parameterization: $\quad \partial c_i D^* \leq \nu_i - 1$

$$\begin{aligned} C_0^* &= K_p^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \\ C^*(s) &= \Lambda(s) - Q(s)N_L(s) = \begin{bmatrix} 2\lambda & 2+\lambda \\ 1-2\lambda & -1-\lambda \end{bmatrix} \qquad \partial r_i C^* = 0 \leq \nu_{max} - 2 \\ D^*(s) &= Q(s)D_L(s) - \Lambda(s)K_p^{-1}H^{-1}(s) \\ &= \begin{bmatrix} -(1+\lambda)s + 7 - 3\lambda & 4-\lambda \\ (2+\lambda)s - 10 + 5\lambda & -6+2\lambda \end{bmatrix} \qquad \begin{array}{l} \partial c_1 D^* = 1 \leq \nu_1 - 1 \\ \partial c_2 D^* = 0 \leq \nu_2 - 1 \end{array} \end{aligned}$$

Assuming that K_p is known, the unknown parameter matrix is given by

$$\theta^{*T} = \begin{bmatrix} 2\lambda & 2+\lambda & \lambda^2-2\lambda+7 & 4-\lambda & -1-\lambda \\ 1-2\lambda & -1-\lambda & -\lambda^2+3\lambda-10 & -6+2\lambda & 2+\lambda \end{bmatrix}$$

where the number of unknown parameter $N_\theta = 10$. The regressor vector is defined by

$$\psi^T = \begin{bmatrix} \frac{u_1}{l(s)(s+\lambda)} & \frac{u_2}{l(s)(s+\lambda)} & \frac{y_{p1}}{l(s)(s+\lambda)} & \frac{y_{p2}}{l(s)(s+\lambda)} & \frac{y_{p1}}{l(s)} \end{bmatrix}$$

Nonidentifiable parameterization: $\qquad \partial c_i D^* \leq \nu - 1 = \nu_{\max} - 1$

$$\begin{aligned}
C_0^* &= K_p^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \\
C^*(s) &= \begin{bmatrix} 2\lambda + k_1 & 2 + \lambda + k_1 \\ 1 - 2\lambda + k_2 & -1 - \lambda + k_2 \end{bmatrix} \qquad \partial r_i C^* = 0 \leq \nu_{\max} - 2 \\
D^*(s) &= \begin{bmatrix} -(1+\lambda+k_1)s + 7 - 3\lambda - 3k_1 & -k_1 s + 4 - \lambda - 2k_1 \\ (2+\lambda-k_2)s - 10 + 5\lambda - 3k_2 & -6 + 2\lambda - 2k_2 \end{bmatrix} \qquad \partial c_i D^* = 1 \leq \nu_{\max} - 1
\end{aligned}$$

where k_1 and k_2 are arbitrary constants. The unknown parameter matrix is given by

$$\theta_n^{*T} = \begin{bmatrix} \theta^{*T} & \begin{matrix} 0 \\ 0 \end{matrix} \end{bmatrix} + \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & \lambda - 3 & \lambda - 2 & -1 & -1 \end{bmatrix}$$

and the regressor vector is defined by

$$\psi_n^T = \begin{bmatrix} \psi^T & \frac{y_{p2}}{l(s)} \end{bmatrix}$$

A simulation of the MRAC algorithm was made for both parameterizations with the following reference input, sufficiently rich to guarantee persistency of excitation (see Theorem 3.1),

$$\begin{aligned}
r_1 &= \sin(2 * t) + \sin(4 * t) \\
r_2 &= \sin(t) + \sin(3 * t)
\end{aligned}$$

and with unmodeled dynamics and high-frequency output noise added to the nominal system. The unmodeled dynamics are $\frac{144}{(s+12)^2}$ on the first output, y_{p_1}, and $\frac{289}{(s+17)^2}$ on the second output, y_{p_2}. The noise is an additive output noise equal to $0.1\sin(20 * t)$ in the first output and equal to $0.1\sin(25 * t)$ in the second output. The unmodeled dynamics and the noise are more than a decade away from the dynamics of the nominal system and of the reference model. Furthermore, the average amplitude of the noise is about 5% of the average amplitude of the reference output. For the simulations, a least-squares algorithm with forgetting factor was used to increase the speed of convergence. Also, a stabilizing term was added to both schemes (*cf.* Kreisselmeier [69]) to prevent divergence of the P matrix (which can occur in the nonidentifiable parameterization). The effect of this term is similar to the resetting but is easier to program. The parameters, λ and l, were selected equal to 4 and 10, and the initial controller parameters were chosen equal to zero. The norm of the output error, $\|e_0\|$, is shown in Fig. 3.1 for the

identifiable parameterization and in Fig. 3.2 for the nonidentifiable parameterization. Clearly, the system becomes unstable with the nonidentifiable parameterization and remains stable with the identifiable one (much longer simulations were run to verify this fact). For illustrative purpose, the estimate parameter C_{11} is shown in Fig. 3.3 in both cases. We can see that the parameter estimates converge around stable tuned values for the parameterization that is identifiable. However, the estimates move out of the stability region for the nonidentifiable parameterization.

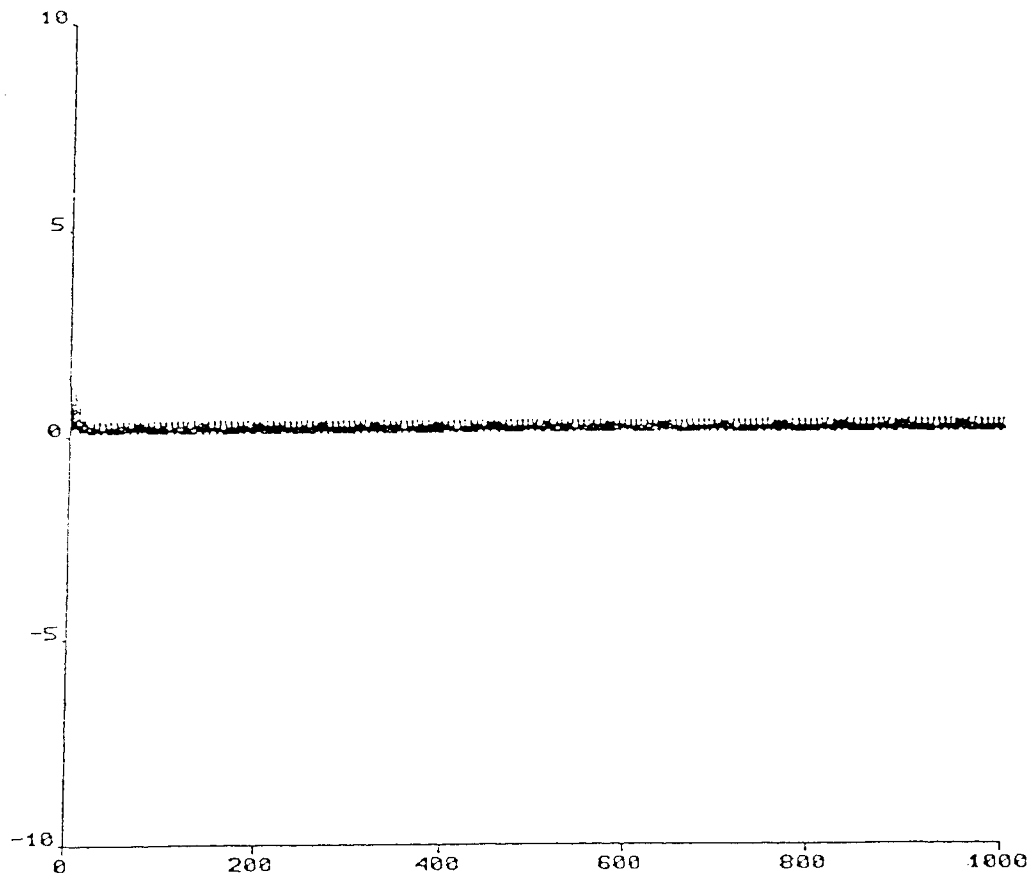

Figure 3.1: $\|e_0(t)\|$ with identifiable parameterization

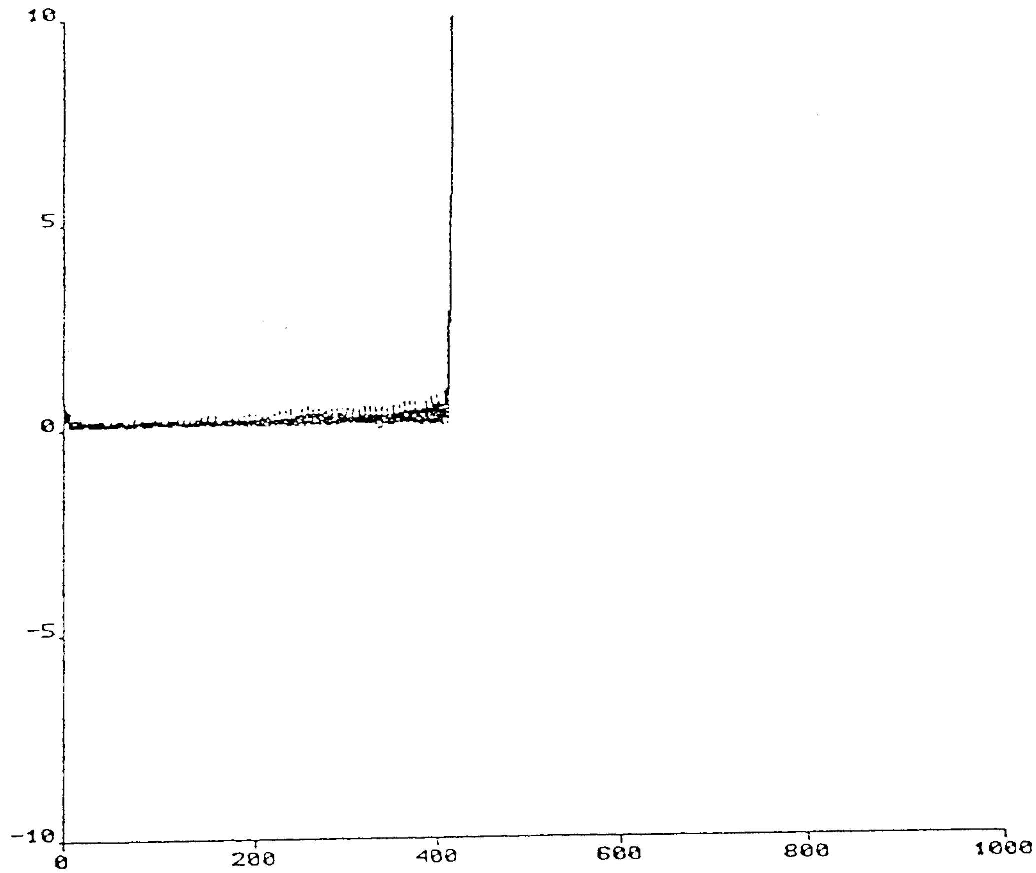

Figure 3.2: $\|e_0(t)\|$ with nonidentifiable parameterization

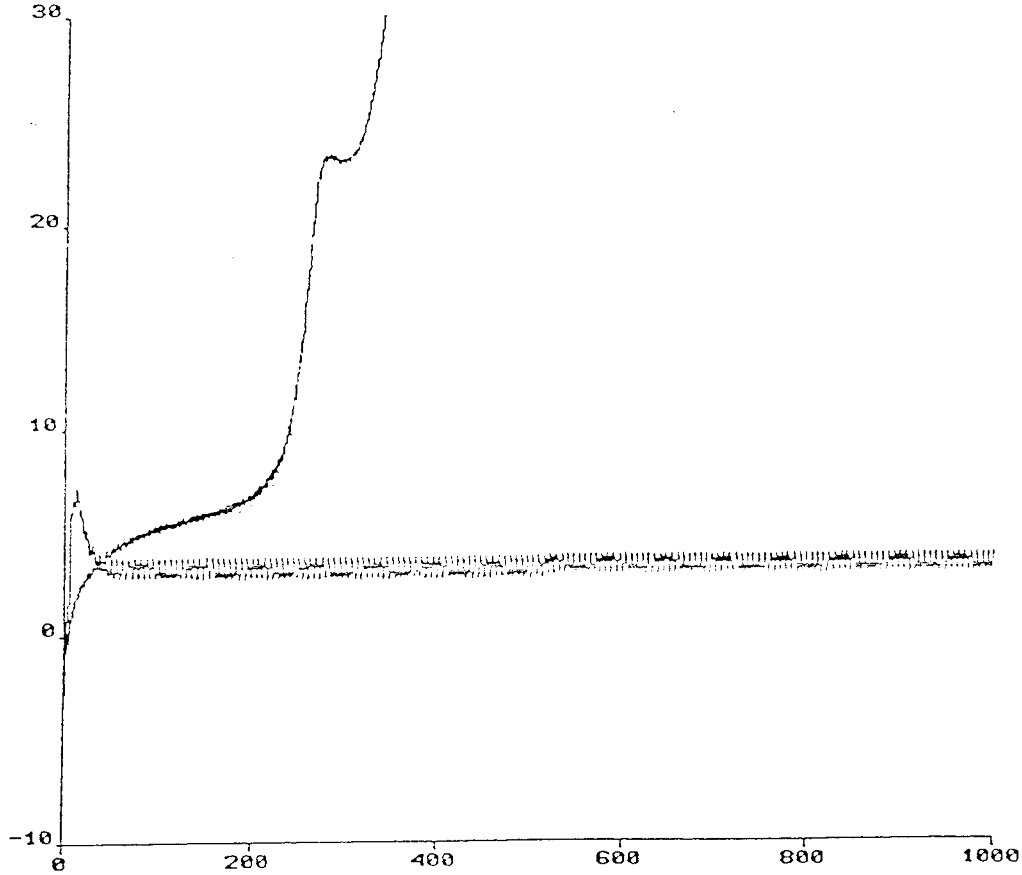

Figure 3.3: $C_{11}(t)$ in both cases

3.4.2 Parameter convergence

The following examples illustrates the results of Table 3.2. Let the plant be a 2-input 2-output system of order 4, with the following transfer function matrix

$$P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

The observability indices are $\nu_1 = 2$ and $\nu_2 = 2$ and the canonical left MFD is

$$N_L(s) = \begin{bmatrix} s & -2s-2 \\ 1 & s-1 \end{bmatrix} \qquad D_L(s) = \begin{bmatrix} s^2+2s+1 & s \\ s+1 & s^2+2s+2 \end{bmatrix}$$

The controllability indices are $\mu_1 = 3$ and $\mu_2 = 1$ and the canonical right MFD is

$$N_R(s) = \begin{bmatrix} s^2+s & -2 \\ 1 & 1 \end{bmatrix} \qquad D_R(s) = \begin{bmatrix} s^3+3s^2+3s+2 & 1 \\ 0 & s+1 \end{bmatrix}$$

Since

$$\begin{aligned} \det N_L(s) = \det N_R(s) &= s^2+s+2 \\ \det D_L(s) = \det D_R(s) &= s^4+4s^3+6s^2+5s+2 = (s+1)(s+2)(s^2+s+1) \end{aligned}$$

the system is stable and minimum phase.

MRAC case – K_p unknown :

Let

$$\Lambda(s) = \begin{bmatrix} (s+\lambda) & 0 \\ 0 & (s+\lambda) \end{bmatrix} \quad \lambda > 0 \qquad L(s) = \begin{bmatrix} (\frac{s}{l}+1) & 0 \\ 0 & (\frac{s}{l}+1) \end{bmatrix} \quad l > 0$$

$$M(s) = H(s) = \begin{bmatrix} \frac{1}{(s+1)} & 0 \\ 0 & \frac{1}{(s+1)} \end{bmatrix}$$

then

$$\begin{aligned} C_0^* &= K_p^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \\ C^*(s) &= \Lambda(s) - Q(s)N_L(s) = \begin{bmatrix} \lambda-2 & 4 \\ -1 & \lambda+1 \end{bmatrix} \qquad \partial r_i C^* = 0 \le \nu_{\max} - 2 \\ D^*(s) &= Q(s)D_L(s) - \Lambda(s)K_p^{-1}H^{-1}(s) \\ &= \begin{bmatrix} (3-\lambda)s+3-\lambda & (3-2\lambda)s+4-2\lambda \\ s+1 & (1-\lambda)s+2-\lambda \end{bmatrix} \qquad \partial c_i D^* = 1 \le \nu_i - 1 \end{aligned}$$

Therefore, the unknown parameter matrix is given by

$$\theta^{*T} = \begin{bmatrix} 1 & 2 & \lambda-2 & 4 & \lambda^2-4\lambda+3 & 2\lambda^2-5\lambda+4 & 3-\lambda & 3-2\lambda \\ 0 & 1 & -1 & \lambda+1 & 1-\lambda & \lambda^2-2\lambda+2 & 1 & 1-\lambda \end{bmatrix}$$

where the number of unknown parameter $N_\theta = 16$. The regressor vector is defined by

$$\psi^T = \begin{bmatrix} \frac{(s+1)y_{P1}}{l(s)} & \frac{(s+1)y_{P2}}{l(s)} & \frac{u_1}{l(s)(s+\lambda)} & \frac{u_2}{l(s)(s+\lambda)} & \frac{y_{P1}}{l(s)(s+\lambda)} & \frac{y_{P2}}{l(s)(s+\lambda)} & \frac{y_{P1}}{l(s)} & \frac{y_{P2}}{l(s)} \end{bmatrix}$$

Finally, the conditions for parameter convergence are

- Sufficient condition: $n + \nu_{\max} - p + 1 = 5$ frequency components per input.
- Necessary condition: $n + p\nu_{\max} = 8$ frequency components in the inputs taken altogether.

PPAC case :

Let

$$\Lambda(s) = \begin{bmatrix} (s+\lambda) & 0 \\ 0 & (s+\lambda)(s+1)^2 \end{bmatrix} \quad \lambda > 0 \qquad L(s) = \begin{bmatrix} (\frac{s}{l}+1) & 0 \\ 0 & (\frac{s}{l}+1) \end{bmatrix} \quad l > 0$$

$$D_M(s) = \begin{bmatrix} (s+1)^2(s+4) & 0 \\ 0 & (s+4) \end{bmatrix} \qquad \partial d_i(s) = \mu_i$$

then

$$\Gamma(s) = \Lambda(s) D_M(s) = \begin{bmatrix} (s+\lambda)(s+1)^2(s+4) & 0 \\ 0 & (s+\lambda)(s+1)^2(s+4) \end{bmatrix}$$

$$\begin{aligned}
V_L^*(s) &= \begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \qquad \partial r_i V_L^* = 0 \le \nu_{\max} - 2 \\
U_L^*(s) &= \begin{bmatrix} 0.5s+1 & 0.5s+2 \\ -0.5s-1 & -0.5s-1 \end{bmatrix} \qquad \partial c_i U_L^* = 1 \le \nu_i - 1 \\
\bar{C}^*(s) &= \Lambda(s) - Q(s)N_L(s) - \Lambda(s)D_M(s)V_L^*(s) = \begin{bmatrix} -3\lambda & 6\lambda-6 \\ 0 & -3s^2-(3\lambda+3)s-3\lambda \end{bmatrix} \\
\bar{C}^*_{0_\infty} &= I - \Gamma_c[D_R]^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{matrix} \partial c_i \bar{C}^* = \{0,2\} \le \{\mu_{\max} - \mu_i + \nu_{\max} - 2\} \\ \partial r_i \bar{C}^* = \{0,2\} \le \{\mu_{\max} - \mu_i + \nu_{\max} - 1\} \end{matrix} \\
\bar{D}^*(s) &= Q(s)D_L(s) - \Lambda(s)D_M(s)U_L^*(s) \\
&= \begin{bmatrix} (3\lambda-3)s+3\lambda-3 & s+4\lambda \\ 0 & 0 \end{bmatrix} \qquad \partial c_i D^* = 1 \le \nu_i - 1
\end{aligned}$$

The unknown parameter matrix is

$$\theta^{*T} = \begin{bmatrix} -3\lambda & 6\lambda-6 & 0 & 0^{(*)} & 3\lambda-3 & 4\lambda & 3\lambda-3 \\ 0 & -3\lambda & -3\lambda-3 & -3 & 0 & 0 & 0 \\ & & 1 & -0.5 & 0.5 & 1 & 2 & 0.5 & 0.5 \\ & & 0 & 0.5 & -0.5 & -1 & -1 & -0.5 & -0.5 \end{bmatrix}$$

where $(*)$ means that the parameter is always equal to zero for all systems of order 4, with observability index 2 and controllability indices $\{3,1\}$. Consequently, $N_\theta = 27$.

$$\psi^T = \left[\frac{u_1}{l(s)\gamma(s)} \ \frac{u_2}{l(s)\gamma(s)} \ \frac{su_2}{l(s)\gamma(s)} \ \frac{s^2u_2}{l(s)\gamma(s)} \ \frac{y_{P1}}{l(s)\gamma(s)} \ \frac{y_{P2}}{l(s)\gamma(s)} \ \frac{sy_{P1}}{l(s)\gamma(s)} \ \frac{sy_{P2}}{l(s)\gamma(s)} \ \frac{u_1}{l(s)} \ \frac{u_2}{l(s)} \ \frac{y_{P1}}{l(s)} \ \frac{y_{P2}}{l(s)} \ \frac{sy_{P1}}{l(s)} \ \frac{sy_{P2}}{l(s)} \right]$$

Finally, the conditions for parameter convergence are

- Sufficient condition: $(m+1)(\nu_{\max}-1) + m\mu_{\max} + \mu_i = 12$ frequency components in input 1 and 10 in input 2.
- Necessary condition: $2m(\nu_{\max}-1) + m\mu_{\max} + n = 14$ frequency components in the inputs taken altogether.

Identification case :

Let

$$L(s) = \begin{bmatrix} s^2 + l_1 s + l_0 & 0 \\ 0 & s^2 + l_1 s + l_0 \end{bmatrix} \qquad l_1, l_0 > 0$$

Then,

$$\theta^{*T} = \begin{bmatrix} 0 & -2 & 1 & -2 & l_0-1 & 0 & l_1-2 & -1 \\ 1 & -1 & 0 & 1 & -1 & l_0-2 & -1 & l_1-2 \end{bmatrix}$$

$$\psi^T = \left[\frac{u_1}{l(s)} \ \frac{u_2}{l(s)} \ \frac{su_1}{l(s)} \ \frac{su_2}{l(s)} \ \frac{y_{P1}}{l(s)} \ \frac{y_{P2}}{l(s)} \ \frac{sy_{P1}}{l(s)} \ \frac{sy_{P2}}{l(s)} \right]$$

Therefore, $N_\theta = 16$. The conditions for parameter convergence are

- Sufficient condition: $n + \nu_{\max} = 6$ frequency components per input.
- Necessary condition: $n + m\nu_{\max} = 8$ frequency components in the inputs taken altogether.

To illustrate the results, it is convenient to compute the smallest eigenvalue of $R_{\psi_m}(0)$ (direct adaptive control) or $R_\psi(0)$ (identification). This number is $\neq 0$ if ψ_m and ψ are PE and $= 0$ otherwise. Table 3 summarizes the results obtained by computing $R_{\psi_m}(0)$ or $R_\psi(0)$ for several inputs. Case number 1, 2, 3, 4, and 5, with 4 frequency components (f.c.) in both inputs shows

Case	Inputs	nbr. f.c.	$\lambda_{min}(R_{\psi_m}(0)$ or $R_\psi(0))$		
			MRAC	PPAC	Ident.
1	$r_1 = \sin(2.0t) + \sin(4.0t)$	4	$\neq 0$	0	$\neq 0$
	$r_2 = \sin(1.0t) + \sin(3.0t)$	4			
2	$r_1 = \sin(2.0t) + \sin(1.19782411t)$	4	0	0	$\neq 0$
	$r_2 = \sin(1.0t) + \sin(3.0t)$	4			
3	$r_1 = \sin(2.0t) + \sin(0.88518347t)$	4	$\neq 0$	0	0
	$r_2 = \sin(1.0t) + \sin(3.0t)$	4			
4	$r_1 = \sin(1.0t) + \sin(3.0t)$	4	0	0	$\neq 0$
	$r_2 = \sin(2.0t) + \sin(1.19782411t)$	4			
5	$r_1 = \sin(1.0t) + \sin(3.0t)$	4	$\neq 0$	0	$\neq 0$
	$r_2 = \sin(2.0t) + \sin(0.88518347t)$	4			
6	$r_1 = \sin(0.57735027t) + \sin(1.53741222t)$	4	$\neq 0$	0	0
	$r_2 = 1.0 + \sin(1.0t) + \sin(3.0t)$	5			
7	$r_1 = \sin(2.0t)$	2	0	0	$\neq 0$
	$r_2 = \sin(1.0t) + \sin(3.0t) + \sin(5.0t)$	6			
8	$r_1 = 1.0 + \sin(2.0t)$	3	0	0	0
	$r_2 = \sin(1.0t) + \sin(3.0t)$	4			
9	$r_1 = \sin(0.5t) + \sin(2.0t) + \sin(4.0t) + \sin(5.0t)$	8	$\neq 0$	$\neq 0$	$\neq 0$
	$r_2 = \sin(0.2t) + \sin(1.0t) + \sin(3.0t)$	6			
10	$r_1 = 1.0 + \sin(0.5t) + \sin(2.0t) + \sin(4.0t)$	7	$\neq 0$	0	$\neq 0$
	$r_2 = \sin(0.2t) + \sin(1.0t) + \sin(3.0t)$	6			
11	$r_1 = \sin(0.5t) + \sin(2.0t) + \sin(4.0t) + \sin(5.0t)$	8	$\neq 0$	0	$\neq 0$
	$r_2 = 1.0 + \sin(1.0t) + \sin(3.0t)$	5			

Table 3.3: Simulation results

that when the necessary condition is respected but the sufficient is not in the MRAC case and in the identification case, parameter convergence will depend on the location of the f.c. in the input. In case 1 there is parameter convergence in the MRAC case and in the identification case. In case 2, the parameters converge to their true values for the identification scheme but not for the MRAC scheme, while in case 3, it is just the opposite. It was found that there exists an infinite number of values for the f.c. such that convergence does not occur, so that those values are not related in any obvious manner to the zeros of the system. Case 4 and case 5 correspond to case 2 and case 3 when the two inputs are permuted. We can see that it has no effect in the MRAC case, but that it changes the convergence properties in the identification case. This is due to the symmetry of the model in the MRAC case. Case 6 with 5 f.c. in input 1 and 4 f.c. in input 2, illustrates also that convergence occurs in some cases and not in others when the sufficient condition is not satisfied. In case 6, there are more f.c. than in case 1, but

the identification scheme will not allow convergence of all the estimates to their true values. Case 7, shows that with a minimum of 2 f.c. in one input, parameter convergence might still occur in the identification case when the necessary condition is respected. On the other hand, case 8 illustrates that 8 f.c. in the whole input are absolutely necessary to have convergence in the identification and the MRAC case. Case $9, 10$ and 11 illustrates that 14 f.c. in the inputs taken altogether is a necessary condition for parameter convergence in the PPAC case.

We also simulated the behavior of the algorithms under these different inputs. To increase the speed of convergence, a least-squares algorithm with forgetting factor was used for the simulation. Fig. 3.4 shows the evolution of the function $v(t) = \ln \|(\theta - \theta^*)\|$ for the direct MRAC scheme in case 1, 2, and 3. Fig. 3.5 shows the evolution of the function $v(t)$ for the recursive identifier in case 1, 2, and 3. Since $\theta \to \theta^*$ exponentially when ψ is PE, the asymptotic behavior of $v(t)$ in the identification case reflects the rate of convergence of the algorithm. When the input is not PE all the estimates do not converge to their true value, and $\lim_{t\to\infty} v(t) = v(\infty) \neq -\infty$.

It is also interesting to see how the output of the system y_p converges toward the reference output y_m. Fig. 3.6 shows the signal $e(t) = \ln \|(y_p - y_m)\|$ with the signal $v(t)$, in case 1, for the MRAC scheme. Note how the output converges at the *same* rate as the parameters.

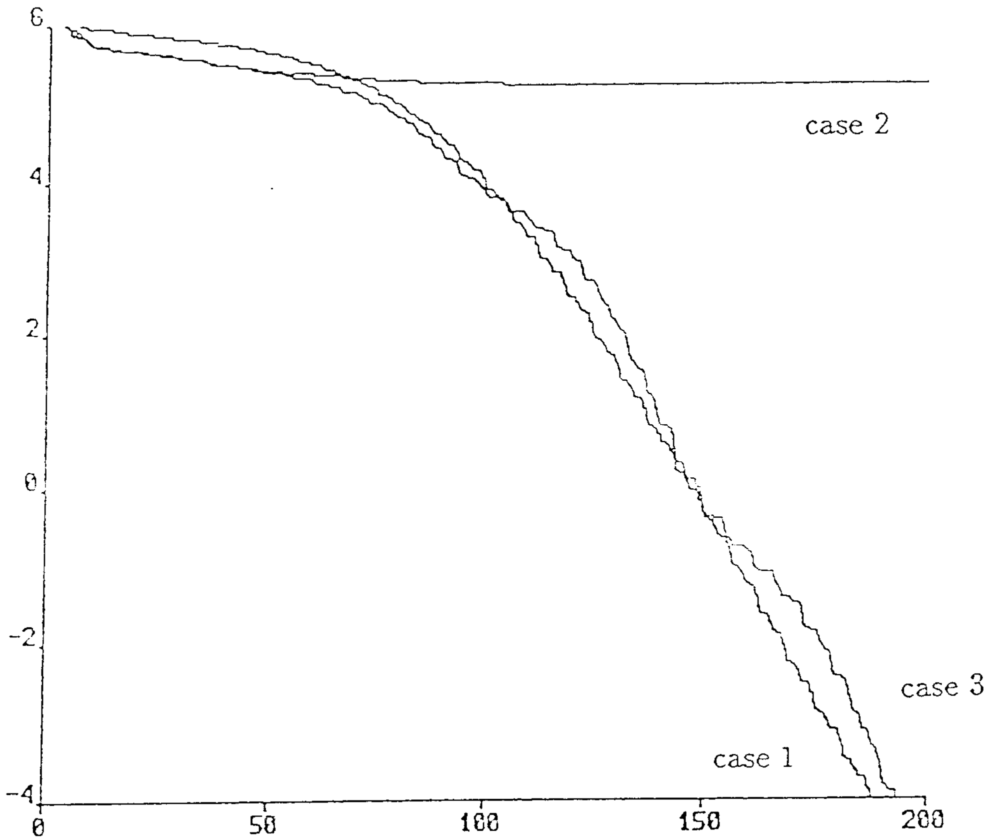

Figure 3.4: Direct MRAC, $v(t)$ in case 1, 2, and 3

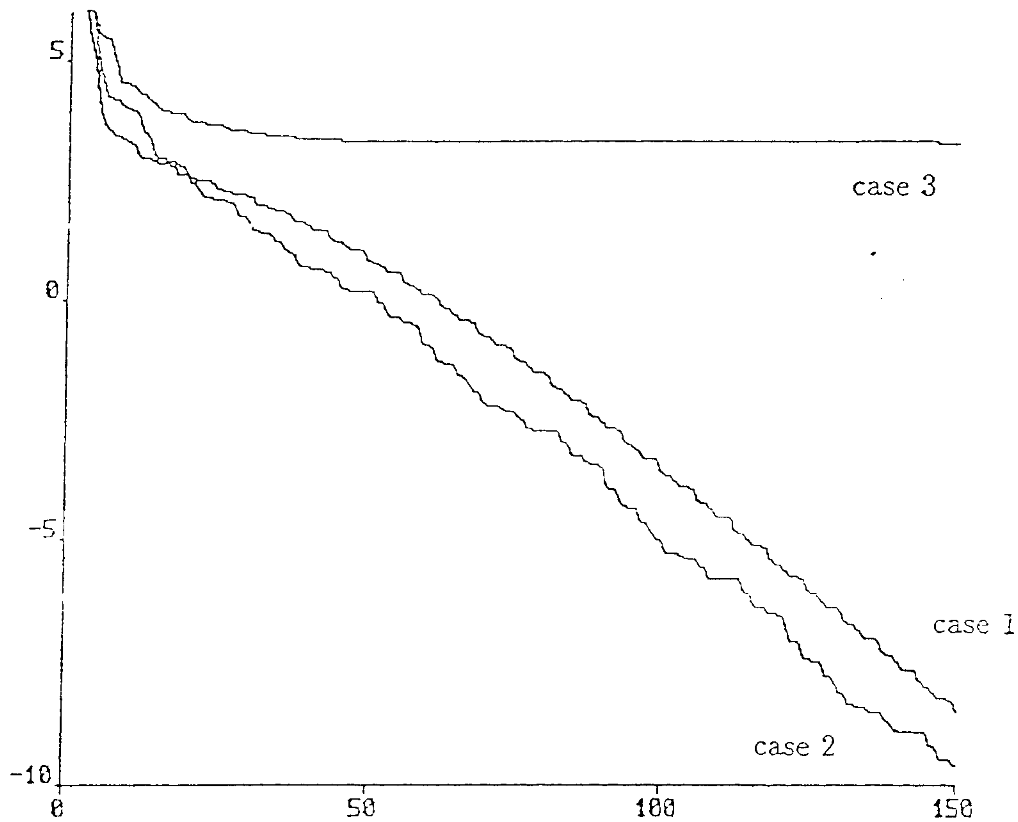

Figure 3.5: Recursive identification, $v(t)$ in case 1, 2, and 3

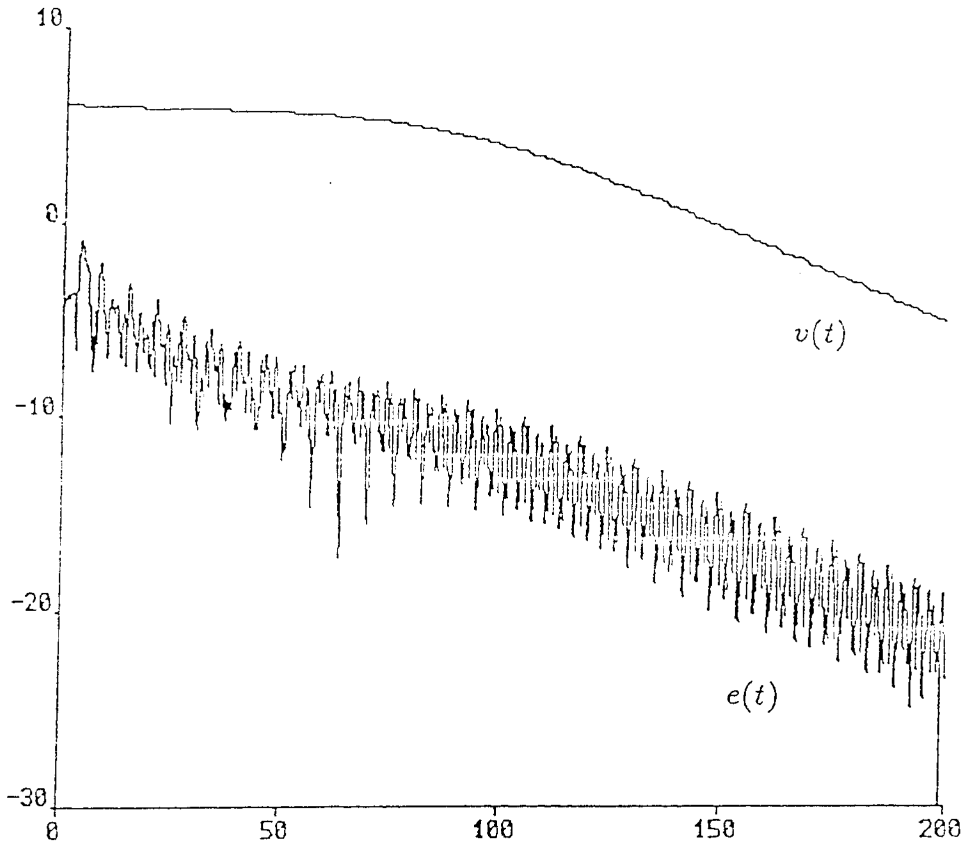

Figure 3.6: Direct MRAC, $e(t)$ and $v(t)$ in case 1

Chapter 4

A transformation to avoid singularity regions

4.0 Introduction

As shown in the previous chapters, SISO adaptive control (MRAC) results have been extended to multivariable systems by several authors, *e.g.*, Goodwin & Long [20], Elliot & Wolovich [6, 11], Singh & Narendra [9, 10], Dugard *et al.* [12], Ortega *et al.* [14], Das [7], Johansson [13], Dion *et al.* [50], and Tao & Ioannou [8]. Unfortunately, similarly to some SISO algorithms requiring the knowledge of the sign of the high-frequency gain, current MIMO algorithms require significant *a priori* knowledge or constraints on the high-frequency gain matrix. Either the high-frequency gain matrix, K_p, must be known (fully or partially, *e.g.*, K_p diagonal with the signs of the diagonal elements known), or it must satisfy some positive definiteness condition (*e.g.*, there exists a known matrix S such that $SK_p + (SK_p)^T > 0$). Another unrealistic alternative is to require that the initial parameter error be sufficiently small so that singularity regions are avoided (*cf.* Theorem 2.2). Note that the discrete-time algorithms presented in Goodwin & Long [20] and in Dugard *et al.* [12] do not explicitly require the knowledge of the time delay gain matrix. In their algorithms the identifier is modified, following Lemma 9.1 in Goodwin *et al.* [19], to avoid singularity of the estimate of the time delay gain matrix. This is necessary because its inverse must be computed in the implementation of the control law (similarly to control law (2.3)). However, even if their modification is implemented, the estimate of the time delay gain matrix may still converge arbitrarily close to singularity if the regressor vector is not sufficiently exciting. Therefore, even if it can be proved in theory that the control signal u is bounded, in practice, u will not be computable when the estimate of the time delay gain matrix converges close to singularity. Furthermore, this modification cannot be adapted to continuous-time systems. Therefore, for all practical purposes, knowledge of the time delay gain matrix or the high-frequency gain matrix must be assumed.

In the SISO case, the problem of relaxing the requirement of knowledge of the sign of the

high-frequency gain was first solved by using controllers based on the so-called Nussbaum gain, *cf.* Mudgett & Morse [70, 71], Morse [72], and Lee & Narendra [73]. Because of the limited practical use of these controllers, Lozano *et al.* [74] proposed a completely different approach. In their algorithm, the controller parameters are obtained from the estimated parameters by applying a transformation with a sort of hysteresis. More specifically, the transformation is the identity when the estimate of the inverse of the high-frequency gain is different from zero. When the estimate becomes too close to zero, the algorithm freezes the controller gain until the estimated gain becomes large enough. Then, the transformation becomes the identity again and the controller parameters are set to be equal to the estimated parameters. A hysteresis guarantees that switching instants are separated by finite time intervals. It is worth noticing that a SISO hysteresis switching controller solving the problem of unknown relative degree and unknown high-frequency gain has also recently been presented by Morse *et al.* [75].

In the MIMO case, the problem of unknown high-frequency gain matrix is even more difficult. The Nussbaum gain approach is not directly applicable to MIMO MRAC algorithms. However, using the hysteresis idea of Lozano *et al.* [74] with some very significant modifications, we show in this chapter how to design a MIMO MRAC algorithm so that stability is guaranteed even if the high-frequency gain matrix is unknown (only an upper bound on the norm of the high-frequency gain matrix and an upper bound on the norm of the matrix of unknown parameters are needed). We show that all the signals in the system remain bounded, that the output error converges to zero, and that the regressor error is in L_2 and converges to zero, independently of the richness of the signals used as reference inputs. We also prove that exponential convergence is achieved when persistency of excitation conditions are met. Aside from the nontrivial multivariable extensions, an original contribution of this chapter are the exponential convergence result (only asymptotic stability is achieved by Lozano *et al.* [74] algorithm). Our success in applying the hysteresis transformation also suggests that such transformation may prove helpful to solve other problems where singularity regions must be avoided, such as in adaptive pole placement, and in nonlinear control using linearization techniques. An adaptive pole-placement SISO algorithm using the hysteresis transformation has just recently been presented by Lozano [76]. Finally, note that an earlier version of the hysteresis transformation was applied to a MIMO MRAC algorithm in de Mathelin & Bodson [77].

4.1 Preliminaries

This section presents definitions and properties from matrix computation theory. While most definitions are standard, many facts reviewed here are not so well-known and will be used in the sequel. For more details about matrix computation theory, the reader is referred to the books by Stewart [78], Stewart & Sun [79], and Golub & Van Loan [80].

Definition 4.1 : Norms

If x is a vector, possibly function of time, then (*cf.* Desoer & Vidyasagar [81], Chapter II), we denote

1. $|x(t)|$ the Euclidean norm of x at time t.
2. $\|x\|_p = [\int_0^\infty |x(\tau)|^p d\tau]^{1/p}$ for $p \in [1, \infty)$
3. $\|x\|_\infty = \sup_{0 \leq t} |x(t)|$
4. $\|x_t\|_p = [\int_0^t |x(\tau)|^p d\tau]^{1/p}$ for $p \in [1, \infty)$
5. $\|x_t\|_\infty = \sup_{0 \leq \tau \leq t} |x(\tau)|$

and we say that $x \in L_p$ when $\|x\|_p$ exists, that $x \in L_{pt}$ when $\|x_t\|_p$ exists for some $t > 0$, and that $x \in L_{pe}$ when $\|x_t\|_p$ exists for all $t < \infty$.

If A is a matrix, possibly function of time, then $\|A\|_F$ will be the Frobenius norm of A and $|A|$, $\|A\|_p$, $\|A\|_\infty$, $\|A_t\|_p$, $\|A_t\|_\infty$, will be the induced norms of A. For example:

$$|A| = \sup_{|x|=1} |Ax|$$

Definition 4.2 : Projections

Following Stewart & Sun [79], Chapter I, let $\mathcal{X}$ be a subspace of $\Re^n$ of dimension $r < n$ and let the columns of the orthogonal matrix $Q_\mathcal{X} \in \Re^{n \times r}$ form an orthonormal basis for $\mathcal{X}$, with $Q_\mathcal{X}^T Q_\mathcal{X} = I$. The matrix

$$P_\mathcal{X} = Q_\mathcal{X} Q_\mathcal{X}^T$$

defines the *orthogonal projection onto* $\mathcal{X}$, *i.e.* $\forall\, x \in \Re^n$, $P_\mathcal{X} x \in \mathcal{X}$ and $(x - P_\mathcal{X} x) \perp \mathcal{X}$. It can be easily verified that the matrix $P_\mathcal{X} \in \Re^{n \times n}$ is symmetric, idempotent, and independent of the choice of $Q_\mathcal{X}$. The matrix

$$P_\mathcal{X}^\perp = I - P_\mathcal{X}$$

defines the *orthogonal projection onto* $\mathcal{X}^\perp$, the subspace of $\Re^n$ of dimension $n - r$, that is the orthogonal complement of $\mathcal{X}$. The matrix $P_\mathcal{X}^\perp \in \Re^{n \times n}$ is also symmetric, idempotent, and independent of the specific choice of $Q_\mathcal{X}$.

Let $\mathcal{R}(A)$ be the *range of* A, *i.e.* the subspace of $\Re^n$ spanned by the columns of the matrix A. Then, there always exists a square orthogonal matrix U such that $AU = \begin{bmatrix} B & 0 \end{bmatrix}$ with B full rank and one has that

$$Q_{\mathcal{R}(A)} = B(B^T B)^{-1/2} \qquad P_{\mathcal{R}(A)} = B(B^T B)^{-1} B^T$$

Definition 4.3 : Singular value decomposition

Let $A \in \Re^{m \times n}$ have rank r, then the *singular value decomposition of* A is given by (cf. Stewart & Sun [79], Chapter I)

$$U^T A V = \Sigma \qquad U^T U = I = V^T V$$

where the matrices $U \in \Re^{m\times \min(m,n)}$, $V \in \Re^{n\times \min(m,n)}$, and $\Sigma = \mathrm{diag}\{\sigma_i\}$ with

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_{\min(m,n)} = 0$$

The matrix of singular values, Σ, is uniquely defined, but the matrices of singular vectors U and V are not. However, if U^* and V^* are another set of such matrices, then it can be shown that there exists an orthogonal matrix $W \in \Re^{\min(m,n)\times\min(m,n)}$, such that

$$U^* = UW \qquad V^* = VW \qquad W\Sigma W^T = \Sigma$$

This implies that UU^T, VV^T and UV^T are invariant. Now, assuming that there are q distinct singular values $\sigma_1^* > \sigma_2^* > \ldots > \sigma_q^* \geq 0$, with multiplicity $\mu_{\sigma_i^*}$ ($\sum_{i=1}^q \mu_{\sigma_i^*} = r$), a further result is that

$$W = \text{block diagonal} \begin{bmatrix} W_1 & & 0 \\ & \ddots & \\ 0 & & W_q \end{bmatrix}$$

where the matrices $W_i \in \Re^{\mu_{\sigma_i^*}\times\mu_{\sigma_i^*}}$ are orthogonal matrices. Then,

$$\begin{bmatrix} U_1 & \ldots & U_q \end{bmatrix}^T A \begin{bmatrix} V_1 & \ldots & V_q \end{bmatrix} = \begin{bmatrix} \Sigma_1 & & 0 \\ & \ddots & \\ 0 & & \Sigma_q \end{bmatrix}$$

is a partioned singular value decomposition of A, where $\Sigma_i = \sigma_i^* I_{\mu_{\sigma_i^*}} \in \Re^{\mu_{\sigma_i^*}\times\mu_{\sigma_i^*}}$ and the products $U_iU_i^T$, $V_iV_i^T$, and $U_iV_i^T$, $i = 1,\ldots,q$, are again unique, despite the fact that the matrices U_i, V_i are not.

Finally, let A and $A+\delta A$ be matrices of same dimensions with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p$ and $\sigma_1 + \delta\sigma_1 \geq \sigma_2 + \delta\sigma_2 \geq \ldots \geq \sigma_p + \delta\sigma_p$. Then, it can be shown (see Stewart & Sun [79], Theorem IV 4.11, pp. 204-205, (by Mirsky)) that

$$\|\mathrm{diag}\,\{\delta\sigma_i\}\,\| \leq \|\delta A\|$$

for any unitarily invariant norm $\|.\|$ (*e.g.*, the Frobenius norm and the induced Euclidean norm). Consequently, the singular values of A are continuous in the elements of A.

Lemma 4.1 : Continuity of subspaces spanned by singular vectors

Let

$$\begin{bmatrix} U_1 & U_2 \end{bmatrix}^T A \begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

be a partitioned singular value decomposition of A, without any constraints on the order in which the singular values appear. Let

$$\begin{bmatrix} U_1 + \delta U_1 & U_2 + \delta U_2 \end{bmatrix}^T (A+\delta A) \begin{bmatrix} V_1 + \delta V_1 & V_2 + \delta V_2 \end{bmatrix} = \begin{bmatrix} \Sigma_1 + \delta\Sigma_1 & 0 \\ 0 & \Sigma_2 + \delta\Sigma_2 \end{bmatrix}$$

be a conformally partitioned singular value decomposition of $A+\delta A$.

1. Suppose that $\exists\, \Delta > 0$ such that $\min(|\sigma_i(\Sigma_1+\delta\Sigma_1)-\sigma_j(\Sigma_2)|) \geq \Delta, \forall\, i, j$, and $\min(\sigma_i(\Sigma_1+\delta\Sigma_1)) \geq \Delta, \forall\, i$. Then

$$(\|P_{\mathcal{R}(U_1)} - P_{\mathcal{R}(U_1+\delta U_1)}\|_F^2 + \|P_{\mathcal{R}(V_1)} - P_{\mathcal{R}(V_1+\delta V_1)}\|_F^2)^{1/2} \leq \frac{2\sqrt{N_1}}{\Delta}\|\delta A\|_F$$

where $N_1 =$ size of Σ_1.

2. Suppose that $\exists\, \Delta, \Gamma > 0$ such that $\min(\sigma_i(\Sigma_1+\delta\Sigma_1)) \geq \Gamma+\Delta, \forall\, i$, and $\max(\sigma_i(\Sigma_2)) \leq \Gamma$, $\forall\, i$. Then

$$\max\{|P_{\mathcal{R}(U_1)} - P_{\mathcal{R}(U_1+\delta U_1)}|, |P_{\mathcal{R}(V_1)} - P_{\mathcal{R}(V_1+\delta V_1)}|\} \leq \frac{|\delta A|}{\Delta}$$

These results, presented under a different form, are due to Wedin and can be found in Stewart & Sun [79] as Theorem V 4.1, pp. 260-262 and Theorem V 4.4, pp. 262. The form presented here is obtained by applying to Wedin's results the properties of the Frobenius norm and Theorem I 5.5, pp. 43-44, in Stewart & Sun [79].

4.2 Hysteresis transformation

Although it might not be obvious *a priori*, the development of a stability proof for the MRAC algorithm reveals that the parameter C_0 in the control law (2.1) must be nonsingular. More precisely, the matrix C_0^{-1} must remain bounded. Since this property is not guaranteed by the parameter identification algorithm (2.29), the parameter C_0 used in the control law will be dissociated from the estimate C_0. Hereafter, we will use the subscript, c, to mark the controller parameters, as opposed to the estimated parameters. So, instead of C_{0_c} being equal to the estimate C_0, a transformation is applied to make sure that C_{0_c} has a bounded inverse even if C_0 is close to singularity. In the SISO case, such a parameter transformation algorithm was introduced by Lozano *et al.* [74]. In this paper, we propose a similar mechanism for the MIMO case which enables us to prove stability. However, some important modifications must be made to the original SISO transformation to make it applicable to the MIMO case.

4.2.1 Assumptions

To implement the parameter transformation, we will assume that an upper bound of the norm of the high-frequency gain matrix, $K_p = C_0^{*-1}$, and that an upper bound of $|\theta^*|$ is known.

4.2.2 Necessary properties

To complete the proof of stability and convergence of the MRAC algorithm, one finds that the following properties are required from the parameter transformation:

1. $C_{0_c}^{-1} \in L_\infty$.
2. $\theta_c \in L_\infty$.
3. $\beta_c = \frac{\phi_c^T \psi}{1+\|\psi_t\|_\infty} \in L_2 \cap L_\infty$, where $\phi_c = \theta_c - \theta^*$.
4. If $\lim_{t\to\infty} \phi(t) = 0$ then $\lim_{t\to\infty} \phi_c(t) = 0$.
5. If $P(t)$ and $\phi(t)$ converge to some P_∞ and ϕ_∞ then $\phi_c(t)$ converges to some $\phi_{c\infty}$.

Since the transformation that we propose to achieve this objective is complex, we proceed to present the transformation in several steps.

4.2.3 Transformation I : simple hysteresis

Define $C_{0_f}(t)$ as being equal to $C_0(t)$ when $\sigma_{\min}(C_0)$ is sufficiently large, and staying constant when $\sigma_{\min}(C_0)$ becomes too small ($\sigma_{\min}$ denotes the smallest singular value). More precisely, since an upper bound on $|K_p|$ is known, a lower bound, σ, on $\sigma_{\min}(C_0^*)$, is also known. Let $\sigma_{\min}(C_0(0)) \geq \sigma$, and define C_{0_f} such that the relationship between $\sigma_{\min}(C_{0_f})$ and $\sigma_{\min}(C_0)$ can be described by Fig. 4.1. Define $\{t_k\}$ as the time instants when the value of C_{0_f} becomes frozen

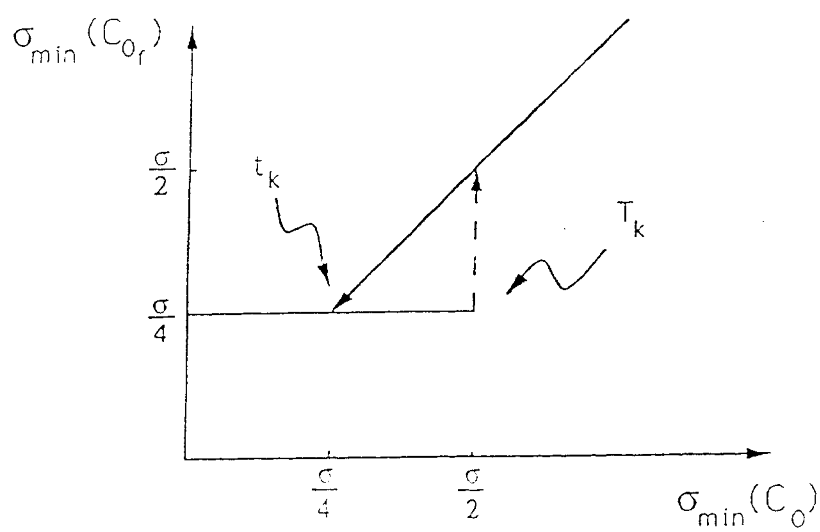

Figure 4.1: $\sigma_{\min}(C_{0_f})$

and $\{T_k\}$ as the time instants when a jump in the value of C_{0_f} occurs (when C_{0_f} is *unfrozen* and becomes equal to C_0 again). The levels at which freezing and unfreezing occur are chosen to be different to prevent repeated switchings arbitrarily closely.

Precisely stated, we let $\sigma_{\min}(C_0(0)) \geq \sigma$, $T_0 = 0$, and $\forall\, k \geq 1$

$$t_k \quad \text{is such that} \quad \begin{cases} \sigma_{\min}(C_0(t_k)) = \sigma/b \\ \sigma_{\min}(C_0(t)) > \sigma/b \quad \forall\, T_{k-1} < t < t_k \end{cases}$$

$$T_k \quad \text{is such that} \quad \begin{cases} \sigma_{\min}(C_0(T_k)) = \sigma/a \\ \sigma_{\min}(C_0(t)) < \sigma/a \quad \forall\, t_k < t < T_k \end{cases} \tag{4.1}$$

In Fig. 4.1, the parameters a and b are respectively equal to 2 and 4, but other values can be selected, provided that $1 < a < b < \infty$.

Note that, since $C_0(t)$ is continuous, $t_k < T_k < t_{k+1} < T_{k+1}\ \forall k \geq 1$ and

$$\begin{aligned} \sigma_{\min}(C_0(t)) &> \sigma/b \quad \text{if} \quad T_{k-1} \leq t < t_k \\ \sigma_{\min}(C_0(t)) &< \sigma/a \quad \text{if} \quad t_k \leq t < T_k \end{aligned}$$

The relationship between $C_{0_f}(t)$ and $C_0(t)$ is described by

$$C_{0_f}(t) = \begin{cases} C_0(t) & \text{if} \quad T_{k-1} \leq t < t_k \\ C_0(t_k) & \text{if} \quad t_k \leq t < T_k \end{cases} \tag{4.2}$$

so that $\sigma_{\min}(C_{0_f}(t)) \geq \sigma/b\ \forall\, t$.

The following parameter transformation

$$\begin{aligned} C_{0_c} &= C_{0_f} \\ C_{i_c} &= C_i \qquad i = 1, \ldots, \nu - 1 \\ D_{i_c} &= D_i \qquad i = 1, \ldots, \nu \end{aligned} \tag{4.3}$$

would then appear to be adequate at this stage. However, it can be checked that the transformation (4.3) does not guarantee property 3 of section 4.2.2, although it guarantees all the others. Therefore, the following additional transformation is necessary.

4.2.4 Transformation II : simple hysteresis + projection

To solve the problem of guaranteeing property 3, the matrices $\{C_{i_c}\}$ and $\{D_{i_c}\}$ in the controller are also transformed, using the following parameter transformation

$$\theta_c(t) = \begin{cases} \theta(t) & \text{if} \quad T_{k-1} \leq t < t_k \quad \forall k > 0 \\ \theta(t) + P(t)P_p(t)(P_p^T(t)P_p(t))^{-1}(C_{0_f}^T(t) - C_0^T(t)) & \text{if} \quad t_k \leq t < T_k \quad \forall k > 0 \end{cases} \tag{4.4}$$

where P_p is the rectangular matrix made of the first p columns of the covariance matrix P (*cf.* Section 2.4). Since P is symmetric, P_p^T is the matrix made of the first p rows of P. It can be

checked that, with this transformation, $C_{0_c} = C_{0_f}$. Further, the transformation is well-defined and all the necessary properties of section 4.2.2 are guaranteed as long as

$$\exists\, \epsilon > 0 \quad \text{s.t.} \quad \sigma_{\min}(P_p^T(t)) > \epsilon \qquad \forall\, t_k \le t < T_k \tag{4.5}$$

In the SISO case, it can be proved that condition (4.5) is always satisfied, so that the parameter transformation (4.4) is always well-defined (it is the transformation presented in Lozano *et al.* [74]). Indeed, given the least-squares estimation algorithm (2.29), $\frac{d}{dt}(P^{-1}\phi) = 0$ between resettings, from which one can deduce that

$$C_0^T(t) = C_0^{*T} + P_p^T(t)\frac{\phi(\tau_j)}{k_0} \tag{4.6}$$

where τ_j is the last covariance resetting time instant, *i.e.* $\tau_j < t < \tau_{j+1}$. Therefore,

$$\sigma_{\min}(C_0^T(t)) \ge \sigma - \sigma_{\max}(P_p^T(t))\frac{|\phi(0)|}{k_0} \tag{4.7}$$

In the SISO case, P_p is a vector. Based on (4.7), if for some $t > 0$, $|P_p(t)| < \sigma(1 - 1/a)\frac{k_0}{|\phi(0)|}$, then the scalar $C_0(t)$ is such that $|C_0(t)| > \sigma/a$ and $\exists\, k \ge 1$ such that $T_{k-1} \le t < t_k$. Therefore, $C_{0_c}(t) = C_0(t)$. In words, this says that if P_p was to be small enough, it would mean that there had been sufficient excitation to identify C_0 accurately enough and the algorithm would be out of the singularity region (first part of (4.4)).

Unfortunately, in the MIMO case, there is no guarantee that condition (4.5) is respected $\forall\, t > 0$. Indeed, if there is excitation in some directions but not others, $\sigma_{\min}(P_p(t))$ could tend to zero while $\sigma_{\max}(P_p(t))$ would not. Therefore, one could conceive that, given certain initial conditions and given certain reference signals, $\lim_{t\to\infty}\sigma_{\min}(P_p(t)) = 0$ with $\lim_{t\to\infty}\sigma_{\min}(C_0(t))$ $< \sigma/b$. In that case, the transformation (4.4) would not be well-defined $\forall\, t > 0$. Therefore, property 2 of section 4.2.2 would not be guaranteed for this transformation in the MIMO case.

4.2.5 Transformation III : selective hysteresis + projection

To solve the problem of the pseudo-inverse $P_p(P_p^T P_p)^{-1}$ not being necessarily bounded in the MIMO case, an additional modification is brought to the parameter transformation. Essentially, the problem is that if the estimated matrix C_0 persists in being singular, despite the presence of excitation in certain (but obviously not all) channels, then the matrix C_{0_c} must be updated selectively in the directions where there is excitation. The major challenge is to achieve this objective while maintaining the boundedness of $C_{0_c}^{-1}$ and the properties listed in section 4.2.2.

We consider the following transformation

$$\theta_c(t) = \begin{cases} \theta(t) & \text{if} \quad T_{k-1} \le t < t_k \quad \forall k > 0 \\ \theta(t) + P(t)V_P(t)\Sigma_P^{-1}(t)U_P^T(t)(C_{0_f}^T(t)R_P^T(t) - C_0^T(t)) & \text{if} \quad t_k \le t < T_k \quad \forall k > 0 \end{cases} \tag{4.8}$$

where $V_P \Sigma_P^{-1} U_P^T$ is a lower rank approximation of $P_p(P_p^T P_p)^{-1}$. Specifically, we consider the following partioned singular value decomposition of P_p^T

$$\left[\begin{array}{cc} U_P & U_\epsilon \end{array}\right]^T P_p^T \left[\begin{array}{cc} V_P & V_\epsilon \end{array}\right] = \left[\begin{array}{cc} \Sigma_P & 0 \\ 0 & \Sigma_\epsilon \end{array}\right] = \Sigma = \operatorname{diag}\{\sigma_i\} \quad \sigma_i \geq \sigma_j \quad \forall\, i < j \tag{4.9}$$

such that

$$\begin{aligned}
\Sigma_P &= \left[\begin{array}{cccc} \sigma_1 & 0 & . & 0 \\ 0 & \sigma_2 & . & . \\ . & . & . & . \\ 0 & . & 0 & \sigma_{N_P} \end{array}\right] > (\sigma_{N_P+1} + \delta_P) I \geq (\epsilon_P + \delta_P) I \\
\Sigma_\epsilon &= \left[\begin{array}{cccc} \sigma_{N_P+1} & 0 & . & 0 \\ . & . & . & . \\ . & . & \sigma_{p-1} & 0 \\ 0 & . & 0 & \sigma_p \end{array}\right] \leq \left[\begin{array}{cccc} \epsilon_P + (p - N_P - 1)\delta_P & 0 & . & 0 \\ . & . & . & . \\ . & . & \epsilon_P + \delta_P & 0 \\ 0 & . & 0 & \epsilon_P \end{array}\right] \\
&\leq (\epsilon_P + (p - N_P - 1)\delta_P) I
\end{aligned} \tag{4.10}$$

N_P is the size of Σ_P and, therefore, $U_P \in \Re^{p \times N_P}$, $V_P \in \Re^{2p\nu \times N_P}$, $U_\epsilon \in \Re^{p \times (p-N_P)}$, and $V_\epsilon \in \Re^{2p\nu \times (p-N_P)}$. The constants ϵ_P and δ_P are arbitrarily chosen, but must satisfy the following conditions

$$\begin{aligned}
0 &< \epsilon_P < (1 - 1/a) \frac{\sigma k_0}{(1 + (p-1)/\alpha_P)(\sigma_{\theta^*} + |\theta(0)|)} \qquad \text{with} \qquad \alpha_P > 1 \\
0 &< \delta_P = \epsilon_P / \alpha_P
\end{aligned} \tag{4.11}$$

where σ_{θ^*} is a known upper bound of $|\theta^*|$. Basically, Σ_ϵ contains the singular values of P_p^T no greater than ϵ_P and the ones which are sufficiently close to them (by increment δ_P). Σ_P contains the larger singular values such that $\sigma_{\min}(\Sigma_P) > \sigma_{\max}(\Sigma_\epsilon) + \delta_P$. The difference δ_P is found to be required to guarantee continuity in the proof.

Finally, $R_P^T = V_f U_0^T$, where V_f is an orthonormal basis of the subspace $\mathcal{X} = \mathcal{R}(C_{0_f} U_P U_P^T)$ of size N_P and U_0 is an orthonormal basis of the subspace $\mathcal{Y} = \mathcal{R}(C_0 U_\epsilon U_\epsilon^T)^\perp$ also of size N_P. Therefore, given the properties of projections (*cf.* Definition 4.2)

$$\begin{aligned}
V_f^T V_f = I \quad &\text{and} \quad V_f V_f^T = P_{\mathcal{R}(C_{0_f} U_P U_P^T)} \triangleq P_\mathcal{X} \in \Re^{p \times p} \\
U_0^T U_0 = I \quad &\text{and} \quad U_0 U_0^T = P^\perp_{\mathcal{R}(C_0 U_\epsilon U_\epsilon^T)} \triangleq P_\mathcal{Y} \in \Re^{p \times p}
\end{aligned} \tag{4.12}$$

If $\sigma_{\min}(P_p^T(t)) > \epsilon_P$ then $N_P = p$, and $P_\mathcal{X} = P_\mathcal{Y} = I$. If $\sigma_{\min}(P_p^T(t)) \leq \epsilon_P$ then $N_P < p$ and $P_\mathcal{X}$ and $P_\mathcal{Y}$ are given by

$$\begin{aligned}
P_\mathcal{X} &= C_{0_f} U_P (U_P^T C_{0_f}^T C_{0_f} U_P)^{-1} U_P^T C_{0_f}^T \\
P_\mathcal{Y} &= (I - C_0 U_\epsilon (U_\epsilon^T C_0^T C_0 U_\epsilon)^{-1} U_\epsilon^T C_0^T)
\end{aligned} \tag{4.13}$$

Indeed, referring to Definition 4.2, we let $A = C_{0_f} U_P U_P^T$ and $U = [\ U_P \quad U_\epsilon\ ]$, so that $B = C_{0_f} U_P$ and $P_{\mathcal{X}}$ is given by (4.13). This assumes that $B = C_{0_f} U_P$ has full column rank, *i.e.*, $U_P^T C_{0_f}^T C_{0_f} U_P$ nonsingular. However, since C_{0_f} is nonsingular, $(U_P^T C_{0_f}^T C_{0_f} U_P)^{-1}$ is well-defined and the number of columns of V_f is equal to N_P. A similar derivation applies for $P_{\mathcal{Y}}$. From (4.6), it can be checked that

$$U_\epsilon^T C_0^T(t) = U_\epsilon^T C_0^{*T} + U_\epsilon^T P_p^T(t) \frac{\phi(\tau_j)}{k_0} = U_\epsilon^T C_0^{*T} + \Sigma_\epsilon V_\epsilon^T \frac{\phi(\tau_j)}{k_0} \tag{4.14}$$

where τ_j is the last covariance resetting time instant, *i.e.* $\tau_j < t < \tau_{j+1}$. Therefore, given (4.10) and (4.11)

$$\sigma_{\min}(U_\epsilon^T C_0^T) \geq \sigma - (\epsilon_P + (p - N_P - 1)\delta_P) \frac{|\phi(0)|}{k_0} > \sigma / a \tag{4.15}$$

so that $(U_\epsilon^T C_0^T C_0 U_\epsilon)^{-1}$ is well-defined and the number of columns of U_0 is indeed equal to N_P.

Comments:

To guarantee some continuity properties and the uniqueness of V_f and U_0, the Gram-Schmidt orthogonalization procedure with memory described in the appendix is applied to $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ to compute V_f and U_0 respectively. Using this procedure, we also have that if $\sigma_{\min}(P_p^T(t)) > \epsilon_P$ then $N_P = p$, $R_P^T(t) = I$, $V_P(t)\Sigma_P^{-1}(t)U_P^T(t) = P_p(t)(P_p^T(t)P_p(t))^{-1}$, and the parameter transformation (4.8) is simply identical to the parameter transformation (4.4).

It can also be shown that $N_P \geq 1$. Indeed, if $\exists\ t > 0$ such that $N_P = 0$, then from (4.15), $\sigma_{\min}(C_0) > \sigma/a$. Therefore, $\exists\ k > 0$ such that $T_{k-1} \leq t < t_k$ and $\theta_c(t) = \theta(t)$. In other words, if P_p was to be small in all directions, then it would mean that there had been enough excitation so that C_0 would be close to C_0^* and one would be out of the singularity region. Therefore, $N_P \geq 1$, $\sigma_{\min}(\Sigma_P) > \epsilon_P + \delta_P$ and the parameter transformation (4.8) is well-defined. Basically, the inverse $(P_p^T P_p)^{-1}$ has been replaced by a lower order inverse, whose existence and boundedness can be guaranteed.

Finally, it can be verified that $C_{0_c}^{-1} \in L_\infty$. Indeed, if $T_{k-1} \leq t < t_k$ or if $t_k \leq t < T_k$ and $\sigma_{\min}(P_p^T(t)) > \epsilon_P$, then $C_{0_c}^T(t) = C_{0_f}^T(t)$ whose inverse is always bounded. If $t_k \leq t < T_k$ and $\sigma_{\min}(P_p^T) \leq \epsilon_P$, then

$$\begin{aligned} C_{0_c}^T(t) &= C_0^T(t) + P_p^T(t) V_P(t) \Sigma_P^{-1}(t) U_P^T(t) (C_{0_f}^T(t) R_P^T(t) - C_0^T(t)) \\ &= C_0^T(t) + U_P(t) U_P^T(t) (C_{0_f}^T(t) R_P^T(t) - C_0^T(t)) \\ &= U_P(t) U_P^T(t) C_{0_f}^T(t) R_P^T(t) + U_\epsilon(t) U_\epsilon^T(t) C_0^T(t) \end{aligned} \tag{4.16}$$

Therefore, using (4.12) and (4.13),

$$\begin{aligned} C_{0_c}^T C_{0_c} &= U_P U_P^T C_{0_f}^T V_f V_f^T C_{0_f} U_P U_P^T + U_\epsilon U_\epsilon^T C_0^T C_0 U_\epsilon U_\epsilon^T \\ &= \begin{bmatrix} U_P & U_\epsilon \end{bmatrix} \begin{bmatrix} U_P^T C_{0_f}^T C_{0_f} U_P & 0 \\ 0 & U_\epsilon^T C_0^T C_0 U_\epsilon \end{bmatrix} \begin{bmatrix} U_P^T \\ U_\epsilon^T \end{bmatrix} \end{aligned} \tag{4.17}$$

so that $\sigma_{\min}(C_{0_c}) \geq \sigma/b$.

4.2.6 Transformation IV : hybrid implementation

In practice $V_P \Sigma_P^{-1} U_P^T$ and R_P^T need not be computed continuously. So, let δ_t be the time interval between two computations and define the time instant t_{0j}^k as the j-th time instant in the time interval $[t_k, T_k)$ where $\sigma_{\min}(P_p^T(t))$ becomes smaller than ϵ_P (because of covariance resetting, there could be several time instants in the time interval $[t_k, T_k)$ where $\sigma_{\min}(P_p^T(t))$ becomes smaller than ϵ_P). Then, define t_{ij}^k as the time instants in the time interval $[t_k, T_k)$ such that $\sigma_{\min}(P_p^T(t_{ij}^k)) \leq \epsilon_P$ and $t_{ij}^k = t_{0j}^k + i\delta_t$. Finally, given (4.8) and all the previous definitions, the parameter transformation with hysteresis will have the following form

$$\theta_c(t) = \theta(t) + P(t)Q(t) \tag{4.18}$$

with

$$Q(t) = \begin{cases} 0 & \text{if} \quad T_{k-1} \leq t < t_k \quad \forall k > 0 \\ P_p(t)(P_p^T(t)P_p(t))^{-1}(C_{0_f}^T(t) - C_0^T(t)) & \text{if} \quad \begin{cases} t_k \leq t < T_k \quad \forall k > 0 \\ \text{and} \quad \sigma_{\min}(P_p^T(t)) > \epsilon_P \end{cases} \\ V_P(t_{ij}^k)\Sigma_P^{-1}(t_{ij}^k)U_P^T(t_{ij}^k)(C_{0_f}^T(t_{ij}^k)R_P^T(t_{ij}^k) - C_0^T(t_{ij}^k)) & \text{if} \quad \begin{cases} t_k \leq t < T_k \quad \forall k > 0 \\ \text{and} \quad \sigma_{\min}(P_p^T(t)) \leq \epsilon_P \end{cases} \end{cases} \tag{4.19}$$

where $t_{ij}^k = t_{0j}^k + i\delta_t \leq t < t_{(i+1)j}^k$ and δ_t is arbitrarily chosen, but satisfies the following condition

$$0 < \delta_t < (1/c)\frac{(\sigma/b)}{k_0(g/\gamma)(|\theta(0)| + \sigma_{\theta^*} + k_0\frac{2(|\theta(0)|+2\sigma_{\theta^*})}{\epsilon_P+\delta_P})} \qquad \text{with} \quad c > 1 \tag{4.20}$$

It can be shown that the parameter transformation (4.18) and (4.19) guarantees all the properties of section 4.2.2. Indeed,

Lemma 4.2 : Properties of the hysteresis transformation
Assuming that $\psi \in L_{\infty e}$ and that the parameter estimation algorithm is defined by (2.29), then

1. The transformation is always uniquely defined.
2. $C_{0_c}^{-1} \in L_\infty$.
3. $Q \in L_\infty$, $\theta_c \in L_\infty$, and $\phi_c \in L_\infty$, where $\phi_c = \theta_c - \theta^*$.
4. $\beta_c = \frac{\phi_c^T \psi}{1+\|\psi_t\|_\infty} \in L_2 \cap L_\infty$.
5. If $\lim_{t\to\infty} \phi(t) = 0$ then $\lim_{t\to\infty} \phi_c(t) = 0$.

6. $(t_k - T_{k-1})$ and $(T_k - t_k)$ are bounded below $\forall k$.

7. $\{T_k\}$ and $\{t_k\}$ are finite sets.

8. $\{t_{0j}^k\}$ is a finite set.

9. If $P(t)$ and $\phi(t)$ converge to some P_∞ and ϕ_∞ then $\phi_c(t)$ converges to some ϕ_{c_∞}.

The proof is in the appendix.

Comments:

The fact that $Q \in L_\infty$ is obtained because, even though $P^{-1}(t)$ is not necessarily bounded as t increases, Σ_P^{-1} exists and is always bounded when $Q \neq 0$. In other words, should P^{-1} be unbounded, either Σ_P^{-1} will continue to exist or the algorithm will come out of the hysteresis region. The fourth property is most important: it shows that the main property on β in the estimation algorithm remains valid when the estimated parameter error ϕ is replaced by the controller parameter error ϕ_c. The fact that the intervals $(t_k - T_{k-1})$ and $(T_k - t_k)$ are bounded below comes from the normalized nature of the adaptation algorithm and from the separation of the freezing and unfreezing levels a and b. It eliminates the possibility of having an infinite number of jumps in a finite interval of time. The seventh property tells us that there is a finite number of passages in the hysteresis loop. After a while, the algorithm will settle and no more jumps will occur. The property is the consequence of the convergence of ϕ (not necessarily to zero). The algorithm could actually settle inside the hysteresis loop. In that event, we will see in the following sections that the stability properties of the adaptive algorithm are preserved. The eighth property comes from the fact that $\{\tau_k\}$ is a finite set when there is not a sufficient excitation. However, $\{t_{ij}^k\}$ could be an infinite set since i is not necessarily finite. This could occur if the algorithm settles inside the hysteresis loop. The last property is obtained because of the difference δ_P between Σ_P and Σ_ϵ and because of the use of the particular Gram-Schmidt orthogonalization procedure defined in appendix.

Note that from (4.18) and (4.19) with (4.16),

$$\begin{aligned} U_P^T(t_{ij}^k)C_{0_c}^T(t_{ij}^k) &= U_P^T(t_{ij}^k)C_{0_f}^T(t_{ij}^k)R_P^T(t_{ij}^k) \\ U_\epsilon^T(t_{ij}^k)C_{0_c}^T(t_{ij}^k) &= U_\epsilon^T(t_{ij}^k)C_0^T(t_{ij}^k) = U_\epsilon^T(t_{ij}^k)C_0^{*T}(t_{ij}^k) + \Sigma_\epsilon(t_{ij}^k)V_\epsilon^T(t_{ij}^k)\frac{\phi(\tau_l)}{k_0} \end{aligned}$$

where τ_l is the last covariance resetting time instant. Basically, this particular parameter transformation unfreezes C_{0_c} in the directions where $(C_0^T - C_0^{*T})$ is sufficiently small, using the knowledge that there has been sufficient excitation in those directions. In other words, the parameter transformation defined by (4.18) and (4.19) checks $\sigma_{\min}(C_0)$, the minimum singular value of C_0. When $\sigma_{\min}(C_0)$ is sufficiently large, $C_{0_c} = C_0$. When $\sigma_{\min}(C_0)$ becomes too small, the value of C_{0_c} is *frozen* (stays constant) until $\sigma_{\min}(C_0)$ becomes sufficiently large again. An important twist, however, is that C_{0_c} is *not* frozen in directions where there is sufficient excitation in the signals.

4.3 Stability analysis

4.3.1 Error formulation

It is useful to represent the adaptive system in terms of its deviation with respect to the ideal situation when $\phi = 0$. Given the definition of the controller parameter matrix θ_c, the control law (2.1) is rewritten as

$$u = C_{0_c} r + \bar{\theta}_c^T \bar{w} = \theta_c^T w \tag{4.21}$$

where $\bar{\theta}_c$ is a submatrix of θ_c, and w is defined as

$$w^T = \left[\begin{array}{cc} r^T & \bar{w}^T \end{array} \right] = \left[\begin{array}{cccc} r^T & w^{(1)^T} & w^{(2)^T} & y_p^T \end{array} \right]$$

with

$$\begin{aligned} w^{(1)^T} &= \left[\begin{array}{ccc} \Lambda^{-1}[u]^T & \dots & s^{(\nu-2)}\Lambda^{-1}[u]^T \end{array} \right] = H_{w^{(1)}u}[u]^T \\ w^{(2)^T} &= \left[\begin{array}{ccc} \Lambda^{-1}[y_p]^T & \dots & s^{(\nu-2)}\Lambda^{-1}[y_p]^T \end{array} \right] = H_{w^{(2)}y_p}[y_p]^T \end{aligned}$$

If we define the signal r_p as

$$r_p = H^{-1}[y_p]$$

then, the matching equality (2.11) applied to u can be rewritten as

$$u = C_0^* r_p + \bar{\theta}^{*^T} \bar{w} = \theta^{*^T} z \tag{4.22}$$

where

$$z^T = \left[\begin{array}{cc} r_p^T & \bar{w}^T \end{array} \right] = l[\psi^T]$$

given the definition of the regressor vector ψ (2.13).

If we define the *control error*, e_c, as

$$e_c = (\theta_c^T - \theta^{*^T}) w = \phi_c^T w$$

Then, from equations (4.21) and (4.22)

$$\begin{aligned} r_p &= C_0^{*^{-1}} C_{0_c} r + C_0^{*^{-1}} \bar{\phi}_c^T \bar{w} = r + C_0^{*^{-1}} \phi_c^T w = r + C_0^{*^{-1}} e_c && (4.23) \\ r &= C_{0_c}^{-1} C_0^* r_p - C_{0_c}^{-1} \bar{\phi}_c^T \bar{w} = r_p - C_{0_c}^{-1} \phi_c^T z && (4.24) \end{aligned}$$

and the *output error*, $e_0 = y_p - y_m$, can be expressed as

$$e_0 = H[r_p - r] = H[C_0^{*^{-1}} \phi_c^T w] = H[C_0^{*^{-1}} e_c] = H[C_{0_c}^{-1} \phi_c^T z] \tag{4.25}$$

Since $y_p = P[u] = H[r_p]$, the control input can also be expressed in terms of the control error as

$$u = P^{-1} H[r_p] = P^{-1} H[r + C_0^{*^{-1}} e_c]$$

Similarly,

$$\bar{w} = \begin{bmatrix} w^{(1)} \\ w^{(2)} \\ y_p \end{bmatrix} = \begin{bmatrix} H_{w^{(1)}u}P^{-1}H \\ H_{w^{(2)}y_p}H \\ H \end{bmatrix} [r + C_0^{*^{-1}} e_c] = H_{\bar{w}r_p}[r + C_0^{*^{-1}} e_c] \tag{4.26}$$

$$\psi = l^{-1}[z] = l^{-1} \begin{bmatrix} I \\ H_{w^{(1)}u}P^{-1}H \\ H_{w^{(2)}y_p}H \\ H \end{bmatrix} [r + C_0^{*^{-1}} e_c] = H_{\psi r_p}[r + C_0^{*^{-1}} e_c] \tag{4.27}$$

where the transfer function matrices $H_{\bar{w}r_p}$, $H_{\psi r_p}$ are both strictly proper and stable. If we define the model signals, ψ_m, as the signals ψ when $\phi = 0$

$$\psi_m = H_{\psi r_p}[r] = H_{\psi_m r}[r] \tag{4.28}$$

then, we can define the *regressor error*, e_ψ, as

$$e_\psi = \psi - \psi_m = H_{\psi_m r}[C_0^{*^{-1}} e_c]$$

4.3.2 Existence and uniqueness

We assume that r is bounded and piecewise continuous. The proof of existence and uniqueness of the solution to the differential equations of the adaptive system is given in the appendix. It consists in first obtaining a state-space description of the overall adaptive system (plant, observers, estimator), so that it can be described by differential equations of the form

$$\dot{x} = f(t, x) \qquad x(t_0) = x_0 \tag{4.29}$$

Then, we use the following lemma.

Lemma 4.3 : Local existence and uniqueness of nonlinear differential equations
Suppose $f(t, x)$ in (4.29) is continuous in t and x $\forall\ t \in [t_0, T]$ and satisfies the following conditions: $\exists$ finite constants $l, h, r > 0$ such that

$$\begin{aligned} |f(t,x) - f(t,y)| &\leq l|x - y| \quad \forall\ x, y \in \mathcal{B}_r(x_0) \quad \forall\ t \in [t_0, T] \\ |f(t, x_0)| &\leq h \quad \forall\ t \in [t_0, T] \end{aligned}$$

where $\mathcal{B}_r(x_0)$ is a ball of radius r centered at x_0.
Then: (4.29) has exactly one solution over the time interval $[t_0, t_0 + \delta]$ whenever

$$h\delta \exp(l\delta) \leq r \quad \text{and} \quad \delta \leq \min((T - t_0), \frac{\rho}{l}, \frac{r}{h + lr})$$

for some constant $0 < \rho < 1$.

See Vidyasagar [82], Theorem 3.4.2, pp. 79-81, for a proof of this result.

Using Lemma 4.3, we show that $\exists\ \delta_0 > 0$ such that the system has a unique solution over $[0, \delta_0]$. Then, we let $x_0 = x(\delta_0)$ and we prove in a similar manner that $\exists\ \delta_1 > 0$ such that the system has a unique solution over $[\delta_0, \delta_0 + \delta_1]$. This operation is repeated to keep extending the solution. Finally, we show that the interval of existence of the solution can be extended indefinitely. Therefore, the system has a unique solution $\forall\ t > 0$ and cannot have a finite escape time ($x \in L_{\infty e}$).

4.3.3 Useful lemmas

The following lemmas are useful to prove the stability of adaptive control schemes.

Lemma 4.4 : Input-output L_p stability

Let $y = H[u]$, where H is a proper, rational, square matrix transfer function. Let h be its impulse response.

If h is stable, then

$\forall p \in [1, \infty)$ and $\forall u \in L_p$

$$\|y\|_p \leq \|h\|_1 \|u\|_p + \|\epsilon\|_p$$

$\forall u \in L_{\infty e}$

$$|y(t)| \leq \|h\|_1 \|u_t\|_\infty + |\epsilon(t)|$$

where $\epsilon(t)$ is an exponentially decreasing term due to initial conditions.

For a proof, *cf.* Desoer & Vidyasagar [81], pp. 241.

Lemma 4.5 : Output-input L_p stability

Let $y = H[u]$, where H is a proper, rational, square matrix transfer function.

If H is minimum phase and if $\exists k_1, k_2 \geq 0\ \forall u, \dot{u} \in L_{pe}$ such that $\|\dot{u}_t\|_p \leq k_1 \|u_t\|_p + k_2$, then

$\exists a_1, a_2 \geq 0$ such that

$$\|u_t\|_p \leq a_1 \|y_t\|_p + a_2$$

Lemma 4.5 is a trivial MIMO extension of Lemma 3.6.2 in Sastry & Bodson [3] when H has a diagonal Hermite form. If H does not have a diagonal Hermite form, then $\exists$ always a stable and minimum phase, proper, rational, diagonal square matrix transfer function L, such that $LH^{-1}[y] = L[u]$ with LH^{-1} stable and proper. Then, the MIMO extension of Lemma 3.6.2 in Sastry & Bodson [3] can be applied to $L[u]$ and the result follows from Lemma 4.4, since LH^{-1} is stable and proper.

Lemma 4.6 : Input-output convergence to zero

Let $y = H[u]$, where H is a stable, rational, square matrix transfer function.

If H is strictly proper and $u \in L_2$ or H is proper and $u \in L_2 \cap L_\infty$ with $\lim_{t\to\infty} u = 0$, then

$y \in L_2 \cap L_\infty$ and $\lim_{t\to\infty} y = 0$.

If H is proper and $u \in L_\infty$ with $\lim_{t\to\infty} u = 0$, then

$y \in L_\infty$ and $\lim_{t\to\infty} y = 0$.

When H is strictly proper, see Desoer & Vidyasagar [81], ex. 5, pp. 242, for the first part of the lemma, and see see Desoer & Vidyasagar [81], Theorem 9, pp. 59, for the second part of the lemma. If H is proper, then $H(s) = H_0 + H_{sp}(s)$ with H_{sp} strictly proper. Therefore, $y = H_0 u + H_{sp}[u]$ and the results follow.

Lemma 4.7 : Output-input convergence to zero

Let $y = H[u]$, where H is a proper, rational, square matrix transfer function.

If H is minimum phase, $\dot{u} \in L_\infty$, $y \in L_\infty$, and $\lim_{t\to\infty} y = 0$, then

$u \in L_\infty$ and $\lim_{t\to\infty} u = 0$.

The proof is in the appendix

4.3.4 Stability proof

The following theorem is the main stability theorem for the MIMO MRAC system. It shows that given any initial condition and any bounded input $r(t)$, the states of the adaptive system remain bounded and the output error converges to zero as $t \to \infty$. Furthermore, the difference between the regressor vector ψ and the corresponding model vector ψ_m converges to zero as $t \to \infty$ and is in L_2.

Theorem 4.1 : Global stability of the direct MRAC system - K_p unknown

Consider the MIMO MRAC system described in Section 2.3.1 Assume that the parameter estimation algorithm (2.29) and the hysteresis transformation (4.18) and (4.19) are used. If the reference input $r \in L_\infty$ and is piecewise continuous, then

- All states of the adaptive system are bounded functions of time.
- The output error $e_0 = y_p - y_m \in L_\infty$ and $\lim_{t\to\infty} e_0 = 0$.
- The regressor error $e_\psi = \psi - \psi_m \in L_\infty \cap L_2$ and $\lim_{t\to\infty} e_\psi = 0$.

For the stability proof, we follow similar steps as for the SISO stability proof of Sastry & Bodson [3]. The significant ingredients of the proof are that there is finite number of passages in the hysteresis loop and that ϕ_c has the same properties as ϕ (*cf.* Lemma 4.2). Hence, the standard stability and convergence properties are preserved.

The layout of the proof is as follows. First, we express properties of the parameter identifier that are independent of the stability of the adaptive control system. Such properties are generally expressed in terms of the identifier error e_2 (*cf.* (2.15)). Then, we show that the main properties are preserved by the parameter transformation. Finally, we transfer the properties of the identifier to the control loop and prove stability.

Proof of Theorem 4.1

A. All signals belong to $L_{\infty e}$

In section 4.3.2, we have shown that $(y_p, w^{(1)}, w^{(2)}, \psi)$ belong to $L_{\infty e}$ and, therefore, $(\bar{w}, w, u, r_p, z)$ belong to $L_{\infty e}$.

B. ψ is regular and all signals are bounded by $\|\psi_t\|_\infty$

Let $x, \dot{x}, \Re_+ \to \Re^n \in L_{\infty e}$. Then, we say x is *regular* iff $\exists k_1, k_2 > 0$ such that

$$|\dot{x}(t)| \leq k_1 \|x_t\|_\infty + k_2$$

Recall from (4.23) that

$$r_p = C_0^{*^{-1}} C_{0_c} r + C_0^{*^{-1}} \bar{\phi}_c^T \bar{w}$$

and from (4.26), that

$$\bar{w} = H_{\bar{w}r_p}[r_p]$$

with $H_{\bar{w}r_p}$ stable and strictly proper. Therefore, applying Lemma 4.4

$$\begin{aligned} |\bar{w}| &\leq k\|(\bar{\phi}_c^T \bar{w})_t\|_\infty + k \\ |\frac{d}{dt}\bar{w}| &\leq k\|(\bar{\phi}_c^T \bar{w})_t\|_\infty + k \end{aligned}$$

for some constant $k > 0$. Again we use the single symbol k, whenever an inequality is valid for some positive constant. Since ϕ_c is bounded

$$|\frac{d}{dt}\bar{w}| \leq k\|\bar{w}_t\|_\infty + k$$

so that $\bar{w}$ is regular. Now, recall from (4.27) that

$$\psi = l^{-1}[z] = \begin{bmatrix} l^{-1}[C_0^{*^{-1}} C_{0_c} r] \\ 0 \end{bmatrix} + \begin{bmatrix} l^{-1}[C_0^{*^{-1}} \bar{\phi}_c^T \bar{w}] \\ l^{-1}[\bar{w}] \end{bmatrix}$$

where $l(s)$ is stable and $\partial l(s) \geq d$. Then, by applying Lemma 4.4

$$\begin{aligned} |\psi| &\leq k\|(\bar{\phi}_c^T \bar{w})_t\|_\infty + k\|\bar{w}_t\|_\infty + k \leq k\|\bar{w}_t\|_\infty + k \\ |\frac{d}{dt}\psi| &\leq k\|(\bar{\phi}_c^T \bar{w})_t\|_\infty + k\|\bar{w}_t\|_\infty + k \leq k\|\bar{w}_t\|_\infty + k \end{aligned}$$

Now, define $\bar{\psi}$ as $l^{-1}[\bar{w}]$. Since $\bar{w}$ is regular, l^{-1} is strictly proper and minimum phase, by applying Lemma 4.5

$$|\bar{w}| \leq k\|\bar{\psi}_t\|_\infty + k$$

and since $|\bar{\psi}| \leq |\psi|$

$$\begin{aligned} |\bar{w}| &\leq k\|\psi_t\|_\infty + k \\ |\frac{d}{dt}\psi| &\leq k\|\psi_t\|_\infty + k \end{aligned}$$

so that ψ is regular. Now, from (4.23)

$$|r_p| \leq k|\bar{w}| + k \leq k\|\psi_t\|_\infty + k$$

and since $z^T = \left[\begin{array}{cc} r_p^T & \bar{w}^T \end{array} \right]$

$$|z| \leq k|\bar{w}| + k \leq k\|\psi_t\|_\infty + k$$

Similarly, since $u = \theta_c^T w$

$$|u| \leq k|\bar{w}| + k \leq k\|\psi_t\|_\infty + k$$

and applying Lemma 4.4 to $y_p = P[u]$

$$|y_p| \leq k\|\psi_t\|_\infty + k$$

Consequently,

$$|u|, |y_p|, |r_p|, |\bar{w}|, |w|, |z| \leq k\|\psi_t\|_\infty + k$$

and if ψ is proved to be bounded, then all signals will be bounded.

C. $\underline{\lim_{t\to\infty} \beta_c = 0}$

This comes from the regularity of ψ. Indeed, let $t_1 \geq t_2 \geq 0$

$$\begin{aligned} &|\beta(t_1) - \beta(t_2)| \\ &= |\frac{\phi^T(t_1)\psi(t_1)}{1 + \|\psi_{t_1}\|_\infty} - |\frac{\phi^T(t_2)\psi(t_2)}{1 + \|\psi_{t_2}\|_\infty}| \\ &= |(\phi^T(t_1) - \phi^T(t_2))\frac{\psi(t_1)}{1 + \|\psi_{t_1}\|_\infty} + \phi^T(t_2)(\frac{\psi(t_1)}{1 + \|\psi_{t_1}\|_\infty} - \frac{\psi(t_2)}{1 + \|\psi_{t_2}\|_\infty})| \\ &\leq |(\phi^T(t_1) - \phi^T(t_2))| + |\phi^T(t_2)||(\frac{(\psi(t_1) - \psi(t_2))}{1 + \|\psi_{t_1}\|_\infty} + \frac{\psi(t_2)}{1 + \|\psi_{t_2}\|_\infty}\frac{(\|\psi_{t_2}\|_\infty - \|\psi_{t_1}\|_\infty)}{1 + \|\psi_{t_1}\|_\infty}| \\ &\leq |t_1 - t_2|\|\dot{\phi}\|_\infty + |t_1 - t_2||\phi(0)|\frac{k\|\psi_{t_1}\|_\infty + k}{1 + \|\psi_{t_1}\|_\infty} + |\phi(0)||\frac{(\|\psi_{t_1}\|_\infty - \|\psi_{t_2}\|_\infty)}{1 + \|\psi_{t_1}\|_\infty}| \\ &\leq |t_1 - t_2|(\|\dot{\phi}\|_\infty + 2k|\phi(0)|) \end{aligned}$$

Therefore, $\forall\, \epsilon > 0,\ \exists\, \delta = \frac{\epsilon}{\|\phi\|_\infty + 2k|\phi(0)|}$, such that $\forall\, t_1, t_2 \geq 0$

$$|t_1 - t_2| < \delta \qquad \Rightarrow \qquad |\beta(t_1) - \beta(t_2)| < \epsilon$$

so that β is uniformly continuous in t. Consequently, $|\beta|^2$ is also uniformly continuous and, since $\beta \in L_2 \cap L_\infty$, by Barbalat's Lemma (*cf.* Lemma 1.2.1 in Sastry & Bodson [3]), $\lim_{t\to\infty} \beta = 0$.

From Lemma 2.3, $\{\tau_k\}$ is a finite set of size N_τ if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE. Therefore, the exact same reasoning can be applied to $\pi = \frac{P\psi}{1+\|\psi_t\|_\infty}$, to show that π is uniformly continuous $\forall\, t > \tau_{N_\tau}$, $\lim_{t\to\infty} \pi = 0$, and $\lim_{t\to\infty} \dot{P} = 0$, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE. Since $\beta_c = \beta + Q^T\pi$ with $Q \in L_\infty$, it follows that $\lim_{t\to\infty} \beta_c = 0$ if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE. Finally, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, then $\lim_{t\to\infty} C_0 = C_0^*$. Therefore, $\exists\, T > 0$ such that $\beta_c(t) = \beta(t)\ \forall\, t > T$, and $\lim_{t\to\infty} \beta_c = 0$.

D. Relationship between output error and identification error

Recall from equation (4.25) that

$$e_0 = y_p - y_m = H[C_{0_c}^{-1}\phi_c^T z]$$

therefore

$$(HL)^{-1}[e_0] = (HL)^{-1}[y_p - y_m] = L^{-1}[C_{0_c}^{-1}\phi_c^T z] \quad \text{with} \quad \psi^T = l^{-1}[z^T] \quad L = \mathrm{diag}\{l\} \qquad (4.30)$$

Now, this can be used to transfer convergence properties from the identifier error to the output error.

E. All signals are bounded, $\lim_{t\to\infty} e_0 = 0$, and $\lim_{t\to\infty} e_\psi = 0$

Suppose that $\{A, B, C\}$ is a minimal realization of $l^{-1}(s)$, *i.e.*, $C(sI - A)^{-1}B = l^{-1}(s)$. Let x and W such that

$$\begin{aligned} \dot{x} &= Ax + Bz^T\phi_c C_{0_c}^{T^{-1}} \\ \dot{W} &= AW + Bz^T \end{aligned}$$

then

$$\begin{aligned} Cx &= l^{-1}[z^T\phi_c C_{0_c}^{T^{-1}}] \\ CW &= l^{-1}[z^T] = \psi^T \end{aligned}$$

so that

$$L^{-1}[C_{0_c}^{-1}\phi_c^T z] = C_{0_c}^{-1}\phi_c^T\psi + (C(x - W\phi_c C_{0_c}^{T^{-1}}))^T \qquad (4.31)$$

Furthermore, $\forall t \notin \mathcal{T} = \{t_k\} \cup \{T_k\} \cup \{\tau_k\} \cup \{t_{ij}^k\}$

$$\frac{d}{dt}(x - W\phi_c C_{0_c}^{T^{-1}}) = A(x - W\phi_c C_{0_c}^{T^{-1}}) - W\frac{d}{dt}(\phi_c C_{0_c}^{T^{-1}}) \qquad (4.32)$$

where $\forall t \notin \mathcal{T}$

$$\frac{d}{dt}(\phi_c C_{0_c}^{T^{-1}}) = \dot{\phi} C_{0_c}^{T^{-1}} + \dot{P} Q C_{0_c}^{T^{-1}} + P \dot{Q} C_{0_c}^{T^{-1}} - \phi_c C_{0_c}^{T^{-1}} \dot{C}_{0_c}^{T} C_{0_c}^{T^{-1}} \tag{4.33}$$

and $\forall t \notin \mathcal{T}$

$$\dot{C}_{0_c}^{T} = \begin{cases} \dot{C}_0^T = -g\frac{P_p^T \psi \psi^T \phi}{1+\gamma\psi^T\psi} & \text{if } T_k < t < t_{k+1} \quad \forall k \\ 0 & \text{if } t_k < t < T_k \text{ and } \sigma_{\min}(P_p^T(t)) > \epsilon_P \quad \forall k \\ \dot{C}_0^T + \dot{P}_p^T Q(t_{ij}^k) = -g\frac{P_p^T \psi \psi^T (\phi + PQ(t_i^k))}{1+\gamma\psi^T\psi} & \text{otherwise} \end{cases}$$

and $\forall t \notin \mathcal{T}$

$$\dot{Q} = \begin{cases} 0 & \text{if } T_k < t < t_{k+1} \\ -g\frac{P\psi\psi^T Q}{1+\gamma\psi^T\psi} + gP_p(P_p^T P_p)^{-1} P_p^T \frac{[P\psi\psi^T + \psi\psi^T P]}{1+\gamma\psi^T\psi} Q - P_p(P_p^T P_p)^{-1} \dot{C}_0^T & \text{if } \begin{cases} t_k < t < T_k \; \forall \, k \\ \sigma_{\min}(P_p^T(t)) > \epsilon_P \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

Define ζ such that

$$\zeta = \begin{cases} g\phi C_0^{T^{-1}} \frac{P_p^T \psi\psi^T \phi}{1+\gamma\psi^T\psi} & \text{if } T_k < t < t_{k+1} \quad \forall k \\ 0 & \text{otherwise} \end{cases}$$

then $\zeta \in L_2 \cap L_\infty$. Define ξ such that

$$\xi = \begin{cases} gP([P_p(P_p^T P_p)^{-1} P_p^T - I]\frac{[P\psi\psi^T + \psi\psi^T P]}{1+\gamma\psi^T\psi} Q + P_p(P_p^T P_p)^{-1} P_p^T \frac{\psi\psi^T\phi}{1+\gamma\psi^T\psi}) & \text{if } \begin{cases} t_k < t < T_k \quad \forall k \\ \sigma_{\min}(P_p^T(t)) > \epsilon_P \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

Then, given the properties of the identifier (*cf.* Lemma 2.3) when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, given the fact that $\exists \; T > 0$ such that $C_{0_c}(t) = C_0(t)$ and $\xi(t) = 0 \; \forall \; t > T$ when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, and given the fact that $(P_p^T P_p)^{-1}$ is bounded if $\sigma_{\min}(P_p^T) > \epsilon_P$, $\xi \in L_2 \cap L_\infty$.

Define $\chi = \sum_i \sum_j \chi_{ij}$ such that $\forall \; i, j$

$$\chi_{ij} = \begin{cases} g\phi_c C_{0_c}^{T^{-1}} \frac{P_p^T \psi\psi^T \phi_c}{1+\gamma\psi^T\psi} - g\frac{P\psi\psi^T PQ}{1+\gamma\psi^T\psi} & \text{if } t_{ij}^k \le t < t_{(i+1)j}^k \quad \forall k \\ 0 & \text{otherwise} \end{cases}$$

Then, given the properties of the identifier (*cf.* Lemma 2.3) when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE and given the fact that $\exists\, T > 0$ such that $C_{0_c}(t) = C_0(t)$ and $\chi(t) = 0\ \forall\, t > T$, when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, $\chi \in L_2 \cap L_\infty$. Given the definition of ζ, ξ, and χ, using (4.33)

$$\frac{d}{dt}(\phi_c C_{0_c}^{T^{-1}}) = (\dot{\phi} + \zeta + \xi + \chi) C_{0_c}^{T^{-1}} \qquad \forall t \notin \mathcal{T} \tag{4.34}$$

Define Δ_{T_k}, Δ_{τ_k}, and $\Delta_{t_{ij}^k}$, as

$$\begin{aligned} \Delta_{T_k} &= (\phi_c C_{0_c}^{T^{-1}})_{|t=T_k} - (\phi_c C_{0_c}^{T^{-1}})_{|t=T_k^-} \\ \Delta_{\tau_k} &= (\phi_c C_{0_c}^{T^{-1}})_{|t=\tau_k} - (\phi_c C_{0_c}^{T^{-1}})_{|t=\tau_k^-} \\ \Delta_{t_{ij}^k} &= (\phi_c C_{0_c}^{T^{-1}})_{|t=t_{ij}^k} - (\phi_c C_{0_c}^{T^{-1}})_{|t=t_{ij}^{k-}} \end{aligned} \tag{4.35}$$

then, Δ_{T_k}, Δ_{τ_k}, and $\Delta_{t_{ij}^k} \in L_\infty$ (*cf.* proof of Lemma 4.2). Also, note that $\Delta_{\tau_k} = 0$ if $\exists\, k^*$ such that $T_{k^*} < \tau_k < t_{k^*+1}$.

Finally, combining together (4.32), (4.34), and (4.35)

$$\begin{aligned} (x - W\phi_c C_{0_c}^{T^{-1}}) &= e^{At}(x - W\phi_c C_{0_c}^{T^{-1}})_{|t=0} \\ &\quad - \int_0^t e^{A(t-\tau)} W(\tau)(\dot{\phi}(\tau) + \zeta(\tau) + \xi(\tau) + \chi(\tau)) C_{0_c}^{T^{-1}}(\tau) d\tau \\ &\quad - \sum_{T_k<t} e^{A(t-T_k)} W(T_k)\Delta_{T_k} s(t-T_k) - \sum_{\tau_k<t} e^{A(t-\tau_k)} W(\tau_k)\Delta_{\tau_k} s(t-\tau_k) \\ &\quad - \sum_{t_{ij}^k<t} e^{A(t-t_{ij}^k)} W(t_{ij}^k)\Delta_{t_{ij}^k} s(t-t_{ij}^k) \end{aligned}$$

where $s(t)$ is the unit step function. Using (4.30) and (4.31)

$$\begin{aligned} (HL)^{-1}[e_0] &= C_{0_c}^{-1}\phi_c^T\psi + (Ce^{At}(x - W\phi_c C_{0_c}^{T^{-1}})_{|t=0})^T \\ &\quad -(\int_0^t Ce^{A(t-\tau)} W(\tau)(\dot{\phi}(\tau) + \zeta(\tau) + \xi(\tau) + \chi(\tau)) C_{0_c}^{T^{-1}} d\tau)^T \\ &\quad -(\sum_{T_k<t} Ce^{A(t-T_k)} W(T_k)\Delta_{T_k} s(t-T_k))^T - (\sum_{\tau_k<t} Ce^{A(t-\tau_k)} W(\tau_k)\Delta_{\tau_k} s(t-\tau_k))^T \\ &\quad -(\sum_{t_{ij}^k<t} Ce^{A(t-t_{ij}^k)} W(t_{ij}^k)\Delta_{t_{ij}^k} s(t-t_{ij}^k))^T \end{aligned}$$

and since, by Lemma 4.4, $|W| \le k\|z_t\|_\infty + |\epsilon(t)| \le k\|\psi_t\|_\infty + k$

$$\begin{aligned} \left|\frac{(HL)^{-1}[e_0]}{1 + \|\psi_t\|_\infty}\right| &\le k|\beta_c| + |\epsilon(t)| + k\int_0^t |e^{A(t-\tau)}||\dot{\phi} + \zeta + \xi + \chi| d\tau + k\sum_{T_k<t} |e^{A(t-T_k)}| s(t-T_k) \\ &\quad + k\sum_{\tau_k<t} |e^{A(t-\tau_k)}||\Delta_{\tau_k}| s(t-\tau_k) + k\sum_{t_{ij}^k<t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}| s(t-t_{ij}^k) \end{aligned}$$

where $\epsilon(t)$ is an exponentially decreasing term. Since $(\dot{\phi}+\zeta+\xi+\chi) \in L_2 \cap L_\infty$, applying Lemma 4.6

$$\begin{aligned} |(HL)^{-1}[e_0]| \;\; &\leq \;\; (\beta^* + k\sum_{T_k<t} |e^{A(t-T_k)}|s(t-T_k) + k\sum_{\tau_k<t} |e^{A(t-\tau_k)}||\Delta_{\tau_k}|s(t-\tau_k) \\ &\quad + k\sum_{t_{ij}^k<t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}|s(t-t_{ij}^k))(1+\|\psi_t\|_\infty) \end{aligned} \tag{4.36}$$

where $\beta^* \in L_2 \cap L_\infty$ and $\lim_{t\to\infty}\beta^* = 0$. Since $(HL)^{-1}(y_p) = L^{-1}[r_p]$, from (4.26) and (4.27)

$$\psi = \begin{bmatrix} I \\ H_{\bar{w}r_p} \end{bmatrix} (HL)^{-1}[y_p]$$

and, applying Lemma 4.4

$$|\psi| \leq k\|(HL)^{-1}[y_p]_t\|_\infty + |\epsilon(t)| \leq k\|(HL)^{-1}[e_0]_t\|_\infty + k \tag{4.37}$$

Therefore, (4.36) is equivalent to

$$\begin{aligned} |(HL)^{-1}[e_0]| \;\; &\leq \;\; (k\beta^* + k\sum_{T_k<t} |e^{A(t-T_k)}|s(t-T_k) + k\sum_{\tau_k<t} |e^{A(t-\tau_k)}||\Delta_{\tau_k}|s(t-\tau_k) \\ &\quad + k\sum_{t_{ij}^k<t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}|s(t-t_{ij}^k))(1+\|(HL)^{-1}[e_0]_t\|_\infty) \end{aligned} \tag{4.38}$$

In this expression, $\lim_{t\to\infty}\beta^* = 0$, $\{T_k\}$ is finite, $\{t_{0j}^k\}$ is finite, and $t_{ij}^k = t_{0j}^k + i\delta_t$, $\forall\, i$. Furthermore, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, $\exists\, k^* \geq 0$ such that $\Delta_{\tau_k} = 0$, $\forall\, k > k^*$, and $\{t_{ij}^k\}$ is a finite set. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, $\{\tau_k\}$ is a finite set, $\lim_{t\to\infty}\phi(t) = \phi_\infty$, $\lim_{t\to\infty}P(t) = P_\infty$, and, from Lemma 4.2, $\lim_{i\to\infty}|\Delta_{t_{ij}^k}| = 0$. Consequently, $\exists T > 0$ such that $\forall t > T$

$$\begin{aligned} k\beta^* + k\sum_{T_k} |e^{A(t-T_k)}|s(t-T_k) + k\sum_{\tau_k} |e^{A(t-\tau_k)}||\Delta_{\tau_k}|s(t-\tau_k) & \\ + k\sum_{t_{ij}^k<t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}|s(t-t_{ij}^k) \;\; &< \;\; 1 \end{aligned}$$

and

$$|(HL)^{-1}[e_0](t)| \leq k\|(HL)^{-1}[e_0]_T\|_\infty + k \qquad \forall\, t > T$$

Note that this can be interpreted as an application of the small gain theorem (*cf.* Sastry & Bodson [3], pp. 149). Since $(HL)^{-1}[e_0] \in L_{\infty e}$, it follows that $(HL)^{-1}[e_0] \in L_\infty$. Then, from (4.38)

$$\begin{aligned} |(HL)^{-1}[e_0]| \;\; &\leq \;\; k\beta^* + k\sum_{T_k<t} |e^{A(t-T_k)}|s(t-T_k) + k\sum_{\tau_k<t} |e^{A(t-\tau_k)}||\Delta_{\tau_k}|s(t-\tau_k) \\ &\quad + k\sum_{t_{ij}^k<t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}|s(t-t_{ij}^k) \end{aligned} \tag{4.39}$$

Therefore, $\lim_{t\to\infty}(HL)^{-1}[e_0] = 0$. From (4.37), it follows that $\psi \in L_\infty$ and, consequently, all signals in the adaptive system are bounded. Furthermore, since

$$(\psi - \psi_m) = \begin{bmatrix} I \\ H_{\bar{w}r_p} \end{bmatrix} (HL)^{-1}[e_0] \tag{4.40}$$

with $H_{\bar{w}r_p}$ strictly proper and stable, by Lemma 4.6, we have that $e_\psi = (\psi - \psi_m) \in L_\infty$ and $\lim_{t\to\infty} e_\psi = 0$. Then, by applying Lemma 4.6 to $(HL)^{-1}[e_0]$, we have that $L^{-1}[e_0] \in L_\infty$ and $\lim_{t\to\infty} L^{-1}[e_0] = 0$. Then, by applying Lemma 4.4 to (4.25)

$$|\frac{d}{dt}e_0| \leq k\|w_t\|_\infty + |\epsilon(t)|$$

so that $\dot{e}_0 \in L_\infty$. Finally, from Lemma 4.7, $e_0 \in L_\infty$ and $\lim_{t\to\infty} e_0 = 0$.

F. $e_\psi \in L_2$

We first show that the error term

$$E(t) = \sum_{t_{ij}^k < t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}|s(t - t_{ij}^k) \in L_2$$

Indeed, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, then $\{t_{ij}^k\}$ is a finite set and $E(t) = k\epsilon(t)$ with $\epsilon(t)$ exponentially decreasing. However, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, $\{t_{ij}^k\}$ could be either a finite set or an infinite set depending on where C_0 and P_p converge. If $\{t_{ij}^k\}$ is a finite set, then again $E(t) = k\epsilon(t)$ with $\epsilon(t)$ exponentially decreasing. Now, suppose that C_0 and P_p converge to some value such that $\{t_{ij}^k\}$ is an infinite set. Since $\{\tau_k\}$ and $\{t_{0j}^k\}$ are finite set, using the fact that $\lim_{i\to\infty}|\Delta_{t_{ij}^k}| = 0$ and the fact that $\lim_{t\to\infty} E(t) = 0$, then $\exists\, T \in \{t_{ij}^k\}$, $T > \max\{t_{0j}^k\} > \max\{\tau_k\}$ such that

$$\begin{aligned}
\int_0^\infty E^2(t)dt &= \int_0^\infty (\sum_{t_{ij}^k<t} |e^{A(t-t_{ij}^k)}||\Delta_{t_{ij}^k}|s(t-t_{ij}^k))^2 dt \\
&\leq \int_0^\infty (k|\epsilon(t)|^2 + k|\epsilon(t)|)dt \\
&\quad + \int_T^\infty (\sum_{l=0}^{\frac{t-T}{\delta_t}-1<n\leq\frac{t-T}{\delta_t}} |e^{A(t-(T+l\delta_t))}||\Delta_{t_{ij}^k=T+l\delta_t}|s(t-(T+l\delta_t)))^2 dt \\
&\leq k + \sum_{s=0}^{\infty} \int_0^{\delta_t} (\sum_{l=0}^{s} |e^{A(t+l\delta_t)}||\Delta_{t_{ij}^k=T+(s-l)\delta_t}|)^2 dt \\
&\leq k + k\sum_{s=0}^{\infty}(\sum_{l=0}^{s} |e^{Al\delta_t}||\Delta_{t_{ij}^k=T+(s-l)\delta_t}|)^2 \\
&\leq k + k(\sum_{s=0}^{\infty}|e^{As\delta_t}|^2)\sum_{l=0}^{\infty}(|\Delta_{t_{ij}^k=T+l\delta_t}|^2 + 2\sum_{v=1}^{\infty}|e^{Av\delta_t}||\Delta_{t_{ij}^k=T+l\delta_t}||\Delta_{t_{ij}^k=T+(l+v)\delta_t}|)
\end{aligned}$$

$$
\begin{aligned}
&\leq \; k + k(\sum_{s=0}^{\infty} |e^{As\delta_t}|^2) \sum_{l=0}^{\infty} (|\Delta_{t_{ij}^k = T + l\delta_t}|^2 + 2(\sum_{v=1}^{\infty} |e^{Av\delta_t}|) |\Delta_{t_{ij}^k = T + l\delta_t}|^2) \\
&\leq \; k + k(\sum_{s=0}^{\infty} |e^{As\delta_t}|^2)(\sum_{l=0}^{\infty} |\Delta_{t_{ij}^k = T + l\delta_t}|^2) \\
&\leq \; k + k \sum_{l=0}^{\infty} |\Delta_{t_{ij}^k = T + l\delta_t}|^2
\end{aligned}
$$

where $\epsilon(t)$ is an exponentially decreasing term. From (7.7), (7.9), and (7.10) in the proof of Lemma 4.2, assuming that T is sufficiently large so that N_P is constant $\forall\, t \geq T$, that the order of selection of the columns in the Gram-Schmidt orthogonalization procedure with memory is constant $\forall\, t \geq T$, and that $|\delta P_p^T| < \min(1, \delta_P^2/4)$ over the time interval $[t, t+\delta_t]$ $\forall\, t \geq T$, then

$$
\begin{aligned}
|\delta U_P \Sigma_P^{-1} V_P^T| &\leq k|\delta P_p^T| + k|\delta P_p^T|^{1/2} \\
|\delta V_f| &\leq k|\delta P_p^T| \\
|\delta U_0| &\leq k|\delta P_p^T| + k|\delta C_0|
\end{aligned}
$$

Therefore, since C_0 and P_p^T are uniformly continuous $\forall\, t \geq T$

$$
\begin{aligned}
\int_0^{\infty} E^2(t)dt \;&\leq\; k + \sum_{l=1}^{\infty}(k|\int_{T+(l-1)\delta_t}^{T+l\delta_t} \dot{P}_p(t)dt|^2 + k|\int_{T+(l-1)\delta_t}^{T+l\delta_t} \dot{P}_p(t)dt| + k|\int_{T+(l-1)\delta_t}^{T+l\delta_t} \dot{C}_0(t)dt|^2) \\
&\leq\; k + k(\sum_{l=1}^{\infty} \int_{T+(l-1)\delta_t}^{T+l\delta_t} |\dot{P}_p(t)|dt)^2 + k\sum_{l=1}^{\infty} \int_{T+(l-1)\delta_t}^{T+l\delta_t} |\dot{P}_p(t)|dt \\
&\quad + k(\sum_{l=1}^{\infty} \int_{T+(l-1)\delta_t}^{T+l\delta_t} |\dot{C}_0(t)|dt)^2 \\
&\leq\; k + k(\int_T^{\infty} |\dot{P}_p(t)|dt)^2 + k\int_T^{\infty} |\dot{P}_p(t)|dt + k(\int_T^{\infty} |\dot{C}_0(t)|dt)^2 \;\leq\; k
\end{aligned}
$$

since $\dot{P}_p$ and $\dot{C}_0 \in L_2 \cap L_1$ when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE (*cf.* Lemma 2.3). Therefore, $E(t) \in L_2$, and, from (4.39), $(HL)^{-1}[e_0] \in L_2$. Then, by applying Lemma 4.4 to (4.40), we have that $e_\psi = (\psi - \psi_m) \in L_2$. □

Finally, note that there are no requirements on the reference signals r except, of course, that they are bounded and piecewise continuous. However, exponential convergence is achieved only if the regressor vector ψ is PE (*cf.* Lemma 2.3). Also, the regressor error signals are proven to be in L_2 (this was not proven for the SISO algorithm of Lozano *et al.* [74]). This result allows to extend the SISO parameter convergence result to this algorithm (see Chapter 3).

4.4 Examples

4.4.1 Example 1

Let the plant be a (2×2) system of order 4, with the following transfer matrix

$$P(s) = \begin{bmatrix} \frac{s^2+s}{s^3+3s^2+3s+2} & \frac{-2s^3-7s^2-7s-4}{s^4+4s^3+6s^2+5s+2} \\ \frac{1}{s^3+3s^2+3s+2} & \frac{s^3+3s^2+3s+1}{s^4+4s^3+6s^2+5s+2} \end{bmatrix}$$

The observability indices are $\nu_1 = 2$ and $\nu_2 = 2$, the system is stable and minimum phase. Let

$$\Lambda(s) = \begin{bmatrix} (s+\lambda) & 0 \\ 0 & (s+\lambda) \end{bmatrix} \quad \lambda > 0 \qquad L(s) = \begin{bmatrix} (\frac{s}{l}+1) & 0 \\ 0 & (\frac{s}{l}+1) \end{bmatrix} \quad l > 0$$

$$M(s) = H(s) = \begin{bmatrix} \frac{1}{(s+1)} & 0 \\ 0 & \frac{1}{(s+1)} \end{bmatrix}$$

then

$$\begin{aligned} C_0^* &= K_p^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} & \sigma_{\min}(C_0^*) = 0.4142 \\ C^*(s) &= \begin{bmatrix} \lambda-2 & 4 \\ -1 & \lambda+1 \end{bmatrix} & \partial r_i C^* = 0 \leq \nu_{\max} - 2 \\ D^*(s) &= \begin{bmatrix} (3-\lambda)s+3-\lambda & (3-2\lambda)s+4-2\lambda \\ s+1 & (1-\lambda)s+2-\lambda \end{bmatrix} & \partial c_i D^* = 1 \leq \nu_i - 1 \end{aligned}$$

Therefore, the unknown parameter matrix is given by

$$\theta^{*T} = \begin{bmatrix} 1 & 2 & \lambda-2 & 4 & \lambda^2-4\lambda+3 & 2\lambda^2-5\lambda+4 & 3-\lambda & 3-2\lambda \\ 0 & 1 & -1 & \lambda+1 & 1-\lambda & \lambda^2-2\lambda+2 & 1 & 1-\lambda \end{bmatrix}$$

where the number of unknown parameter $N_\theta = 16$. The regressor vector is defined by

$$\psi^T = \begin{bmatrix} \frac{(s+1)y_{P_1}}{l(s)} & \frac{(s+1)y_{P_2}}{l(s)} & \frac{u_1}{l(s)(s+\lambda)} & \frac{u_2}{l(s)(s+\lambda)} & \frac{y_{P_1}}{l(s)(s+\lambda)} & \frac{y_{P_2}}{l(s)(s+\lambda)} & \frac{y_{P_1}}{l(s)} & \frac{y_{P_2}}{l(s)} \end{bmatrix}$$

We present simulations of the adaptive system with and without the hysteresis transformation to show how important the parameter transformation algorithm is for convergence and stability. In our simulations, the different parameters are set to the following values: $\lambda = 4$, $l = 10$, $g = 10$, $\gamma = 0$, $k_0 = 1$, $k_1 = 0$, $k_2 = 0$, $\sigma = 0.1$, $a = 2$, $b = 4$, $\epsilon_P = 10^{-6}$, $\alpha_P = 2$, and the initial estimate of the parameters

$$\theta^T(0) = \begin{bmatrix} -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore, $\sigma_{\min}(C_0(0)) = 0.5 > \sigma$. The reference inputs

$$\begin{aligned} r_1 &= \sin(5t) + \sin(7t) + \sin(10t) \\ r_2 &= \sin(6t) + \sin(8t) + \sin(9t) \end{aligned}$$

contain 6 frequency components each and, consequently, are sufficiently rich to guarantee persistency of excitation (see Chapter 3, Theorem 3.2).

In the first simulation, the adaptive control algorithm has been implemented with the hysteresis transformation. The evolution of the norm of the output error, $e_0 = y_p - y_m$, is shown in Fig. 4.2. After a transient of about 20 seconds, the output error rapidly converges to zero. The singular values $\sigma_{\min}(C_{0_c})$ and $\sigma_{\min}(C_0)$ are compared in Fig. 4.3 and Fig. 4.4 for, respectively, the first 5 and 20 seconds of the simulation. One sees that the system goes three times into the hysteresis before finally settling outside the hysteresis region. The time instants (in seconds) for entering and leaving the hysteresis are $t_1 = 0.165$, $T_1 = 0.208$, $t_2 = 0.509$, $T_2 = 1.762$, $t_3 = 2.224$, and $T_3 = 16.319$. It can also be seen that $\sigma_{\min}(C_0)$ goes to zero 6 times during the simulation before converging toward $\sigma_{\min}(C_0^*) = 0.4142$ (see Fig. 4.5). During the simulation, $\sigma_{\min}(P_p^T)$ stayed above ϵ_P inside the hysteresis region. For illustrative purposes, the controller parameters $C_{0_{c_{11}}}$ and $C_{0_{c_{22}}}$ are shown in Fig. 4.6 and Fig. 4.7 for, respectively, the first 5 and 50 seconds of the simulation. In our simulations, we observed that generally one passage or no passage at all into the hysteresis are more likely than multiple passages. To obtain multiple passages in the simulation, we selected input frequencies well above the bandwidth of the reference model (which would not typically happen in an adaptive control system).

Finally, the same simulation was executed without the parameter transformation algorithm. The norm of the output error is shown in Fig. 4.8 and $\sigma_{\min}(C_0)$ in Fig. 4.9. Clearly, after a first quick passage through the zero region, $\sigma_{\min}(C_0)$ converges into the zero region and does not leave. The control parameter matrix $C_{0_c} = C_0$ settles in a singularity region and, consequently, suppresses part of the reference input excitation. Model matching is now impossible and, as Fig. 4.8 clearly shows, the output error does not converge to zero. These simulations demonstrate how important the transformation is: without the transformation, convergence might not occur.

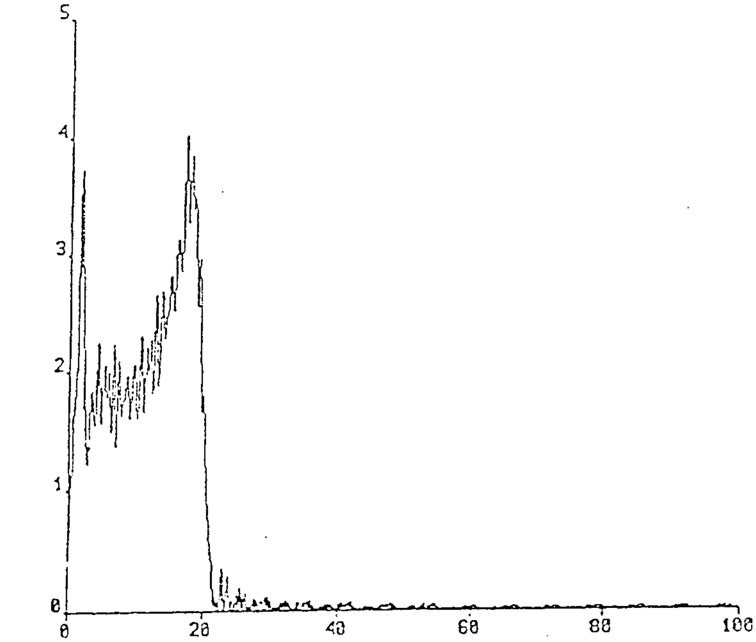

Figure 4.2: $|e_0(t)|$

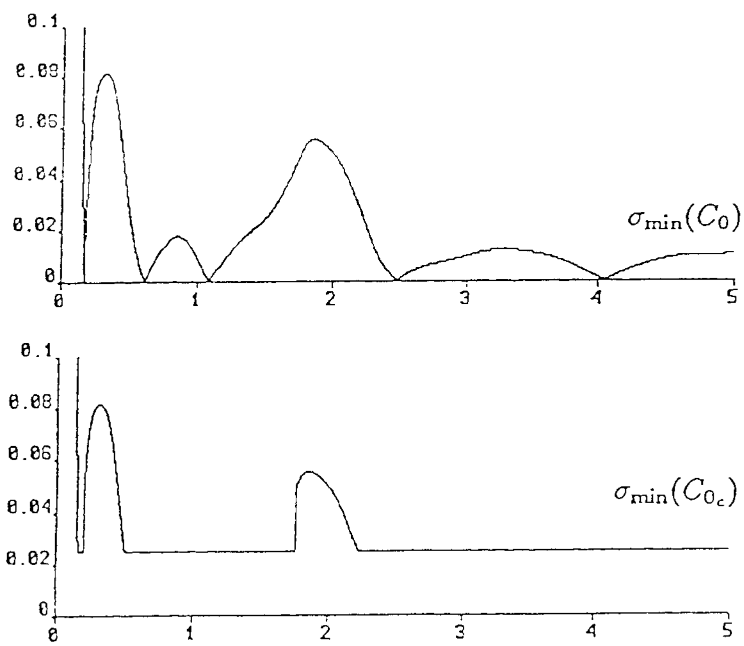

Figure 4.3: $\sigma_{\min}(C_0(t))$ and $\sigma_{\min}(C_{0_c}(t))$

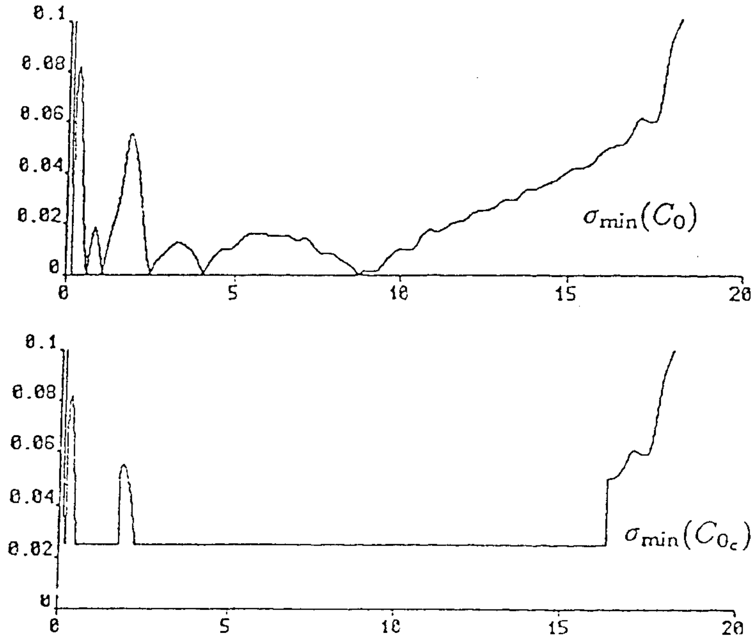

Figure 4.4: $\sigma_{\min}(C_0(t))$ and $\sigma_{\min}(C_{0_c}(t))$

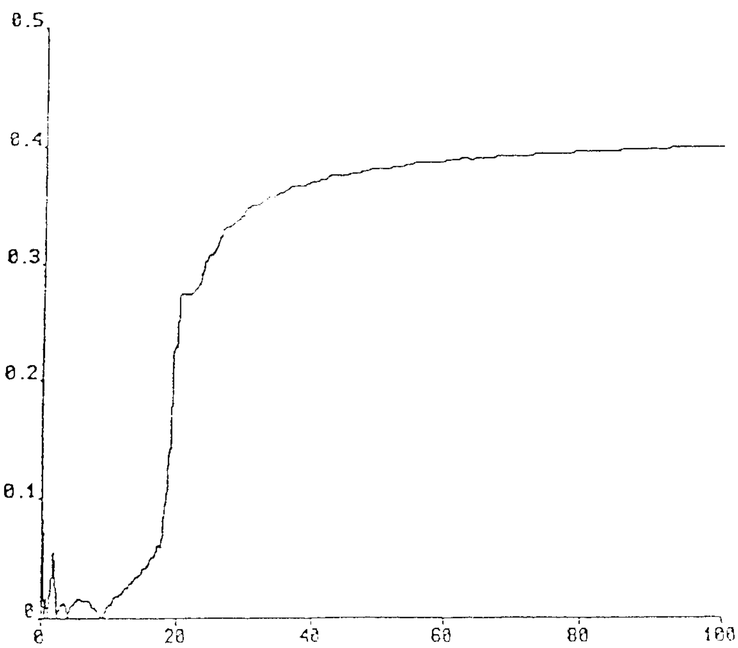

Figure 4.5: $\sigma_{\min}(C_0(t))$

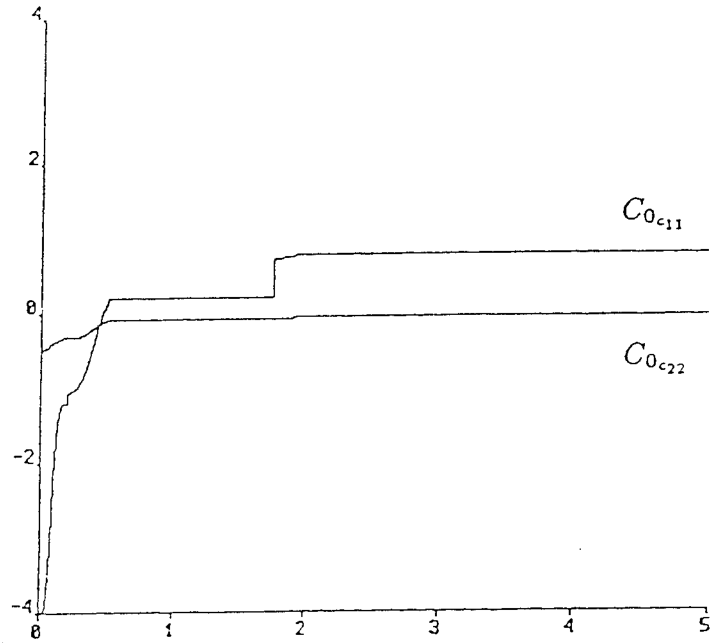

Figure 4.6: $C_{0_{c_{11}}}(t)$ and $C_{0_{c_{22}}}(t)$

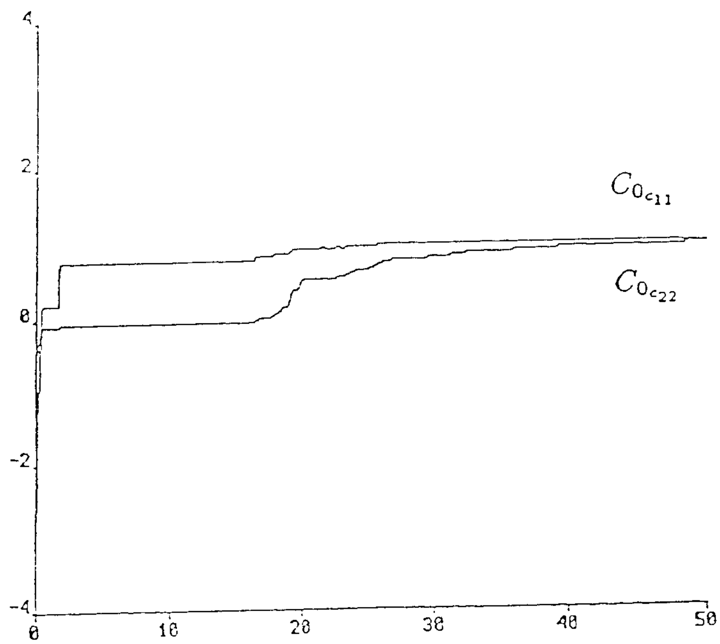

Figure 4.7: $C_{0_{c_{11}}}(t)$ and $C_{0_{c_{22}}}(t)$

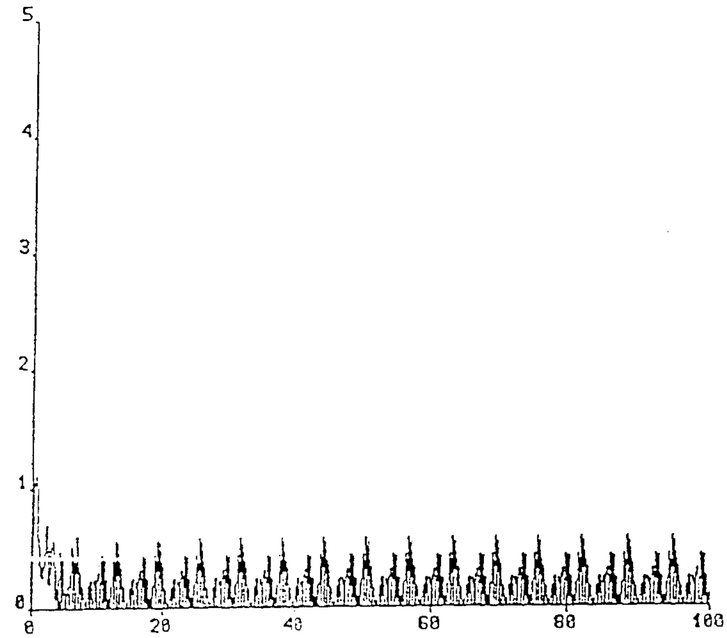

Figure 4.8: $|e_0(t)|$ (without parameter transformation)

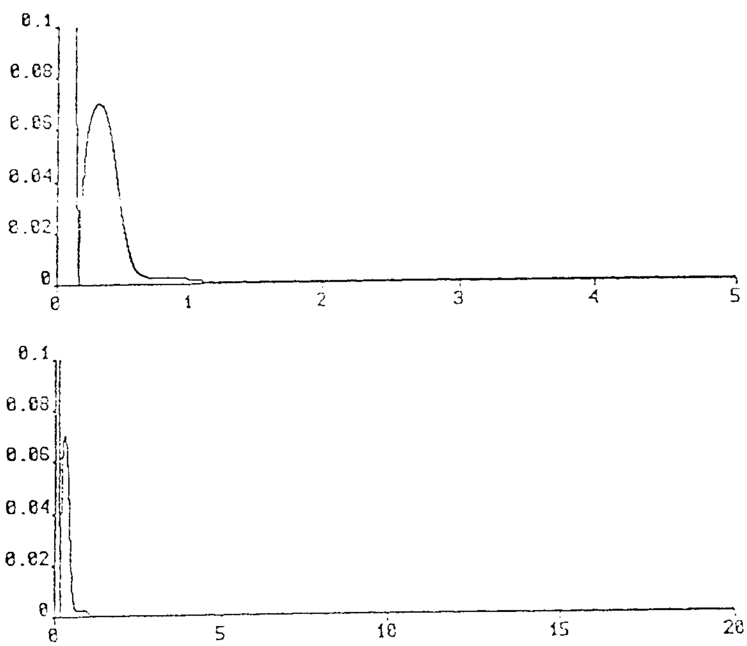

Figure 4.9: $\sigma_{\min}(C_0(t))$ (without parameter transformation)

4.4.2 Example 2

Let the plant be a (2×2) system of order 3, with the following transfer matrix

$$P(s) = \begin{bmatrix} \frac{s+1}{s^2+s+4} & 0 \\ 0 & \frac{1}{s+2} \end{bmatrix}$$

The observability indices are $\nu_1 = 2$ and $\nu_2 = 1$, the system is stable and minimum phase. Let $\Lambda(s)$, $L(s)$, and $M(s) = H(s)$ be as in Example 1, then

$$\begin{aligned} C_0^* &= K_p^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \sigma_{\min}(C_0^*) = 1 \\ C^*(s) &= \begin{bmatrix} \lambda - 1 & 0 \\ 0 & 1 \end{bmatrix} & \partial r_i C^* = 0 \leq \nu_{\max} - 2 \\ D^*(s) &= \begin{bmatrix} -\lambda s + 4 - \lambda & 0 \\ 0 & \lambda - 2 \end{bmatrix} & \begin{matrix} \partial c_1 D^* = 1 \leq \nu_1 - 1 \\ \partial c_2 D^* = 0 \leq \nu_2 - 1 \end{matrix} \end{aligned}$$

Therefore, the unknown parameter matrix is given by

$$\theta^{*T} = \begin{bmatrix} 1 & 0 & \lambda - 1 & 0 & \lambda^2 - \lambda + 4 & 0 & -\lambda \\ 0 & 1 & 0 & 1 & 0 & \lambda - 2 & 0 \end{bmatrix}$$

where the number of unknown parameter $N_\theta = 14$. The regressor vector is defined by

$$\psi^T = \begin{bmatrix} \frac{(s+1)y_{p_1}}{l(s)} & \frac{(s+1)y_{p_2}}{l(s)} & \frac{u_1}{l(s)(s+\lambda)} & \frac{u_2}{l(s)(s+\lambda)} & \frac{y_{p_1}}{l(s)(s+\lambda)} & \frac{y_{p_2}}{l(s)(s+\lambda)} & \frac{y_{p_1}}{l(s)} \end{bmatrix}$$

We present simulations of this decoupled multivariable adaptive system with $\epsilon_P > 0$ (selective hysteresis + projection) and with $\epsilon_P = 0$ (simple hysteresis + projection) to illustrate the necessity of the selective hysteresis. In our simulations, the different parameters are set to the following values: $\lambda = 2$, $l = 10$, $g = 20$, $\gamma = 0$, $k_0 = 1$, $k_1 = 0$, $k_2 = 0$, $\sigma = 0.8$, $a = 5$, $b = 10$, $\alpha_P = 100$, $\delta_t = 0$ (for a 2 input 2 output system, all the quantities in the parameter transformation have an analytic expression and the hybrid implementation was not implemented in the simulation), and the initial estimate of the parameters

$$\theta^T(0) = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The reference inputs are

$$\begin{aligned} r_1 &= \begin{cases} \sin(2t) + \cos(4t) & \text{if } t < 4 \\ 0 & \text{if } t \geq 4 \end{cases} \\ r_2 &= \begin{cases} 0 & \text{if } t < 12 \\ \sin(t) + \sin(3t) & \text{if } t \geq 12 \end{cases} \end{aligned}$$

so that r_1 contains some excitation only $\forall\, t < 4$ and r_2 only $\forall\, t \geq 12$.

In the first simulation, the adaptive control algorithm has been implemented with $\epsilon_P = 0.08$. The evolution of the norm of the output error, $e_0 = y_p - y_m$, is shown in Fig. 4.10. The singular values $\sigma_{\min}(C_{0_c})$ and $\sigma_{\min}(C_0)$ are compared in Fig. 4.11, the parameters $C_{0_{c11}}$ and $C_{0_{11}}$ in Fig. 4.12, and the parameters $C_{0_{c22}}$ and $C_{0_{22}}$ in Fig. 4.13. At first, there is only excitation in the first input, $C_{0_{11}}(t)$ starts converging from -1 to 1, and $C_{0_{22}}(t)$ stays equal to -1. At $t_1 = 0.22$, the system enters the hysteresis region when $\sigma_{\min}(C_0)(t_1) = |C_{0_{11}}(t_1)| = \sigma/b = 0.08$ and $C_{0_c}(t)$ is frozen at its value at t_1 ($C_0(t_1)$). After $t = 4$, there is no signal in the first input any more and the coefficient $C_{0_{11}}(t)$ converges toward 0.113, inside the hysteresis region ($0.113 < \sigma/a = 0.16$). In the mean time, $C_{0_c}(t)$ stays frozen at its value at t_1 (*i.e.*, $C_{0_{c11}} = -0.08$, $C_{0_{c22}} = -1$, and the off-diagonal terms are zero). After $t = 12$, there is excitation in the second input and $C_{0_{22}}$ starts converging from -1 to 1. However, C_{0_c} is still frozen. At $t_{01}^1 = 12.42$, $\sigma_{\min}(P_p^T)$ becomes smaller than $\epsilon_P = 0.08$ and C_{0_c} is partially unfrozen (*i.e.*, $C_{0_{c22}}$ is unfrozen and $C_{0_{c11}}$ stays frozen). Finally, after a short transient, the output error converges to zero. Note that $\sigma_{\min}(C_0(t))$ going to zero at $t = 12.12$ in Fig. 4.11 corresponds to $C_{0_{22}}(t)$ passing through zero.

The same simulation was executed with the simple hysteresis + projection ($\epsilon_P = 0$) instead of the selective hysteresis + projection ($\epsilon_P > 0$). The norm of the output error is shown in Fig. 4.14. After $t_1 = 0.22$, C_{0_c} becomes frozen. However, C_{0_c} stays completely frozen even when there is excitation in the second input ($t > 12$) and the signals rapidly become unbounded. These simulations clearly illustrate the importance of the selective hysteresis.

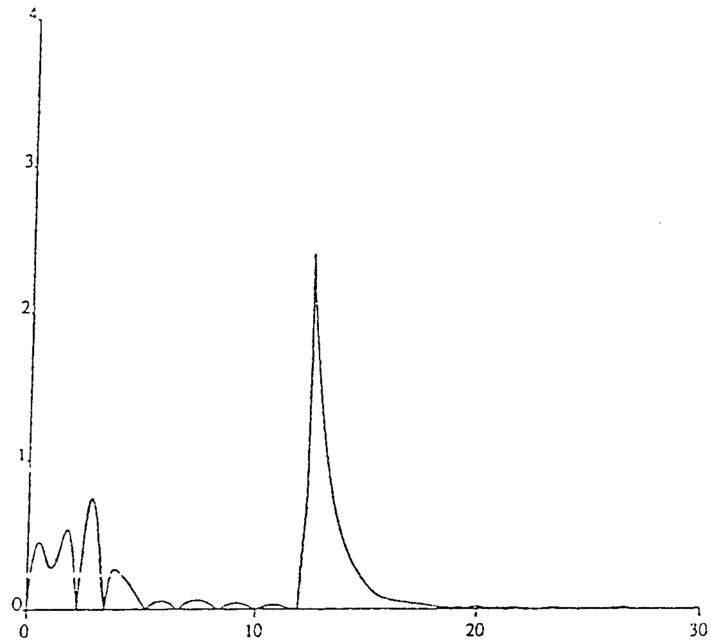

Figure 4.10: $|e_0(t)|$ (with $\epsilon_P = 0.08$)

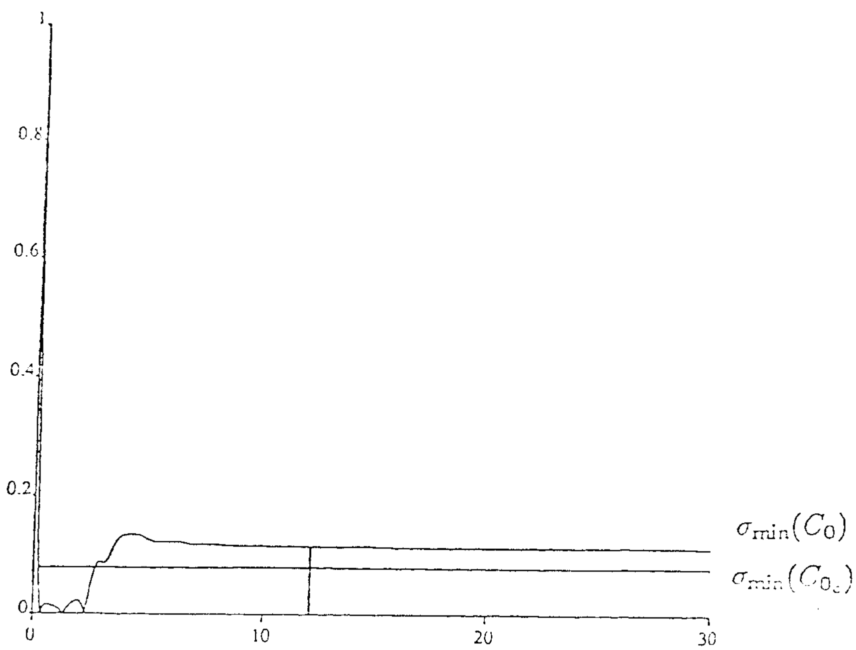

Figure 4.11: $\sigma_{\min}(C_0(t))$ and $\sigma_{\min}(C_{0_c}(t))$ (with $\epsilon_P = 0.08$)

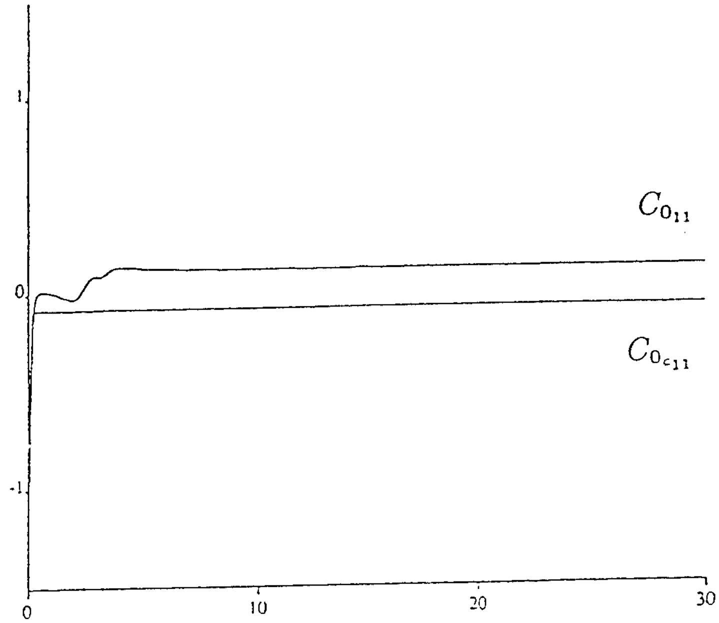

Figure 4.12: $C_{0_{c_{11}}}(t)$ and $C_{0_{11}}(t)$ (with $\epsilon_P = 0.08$)

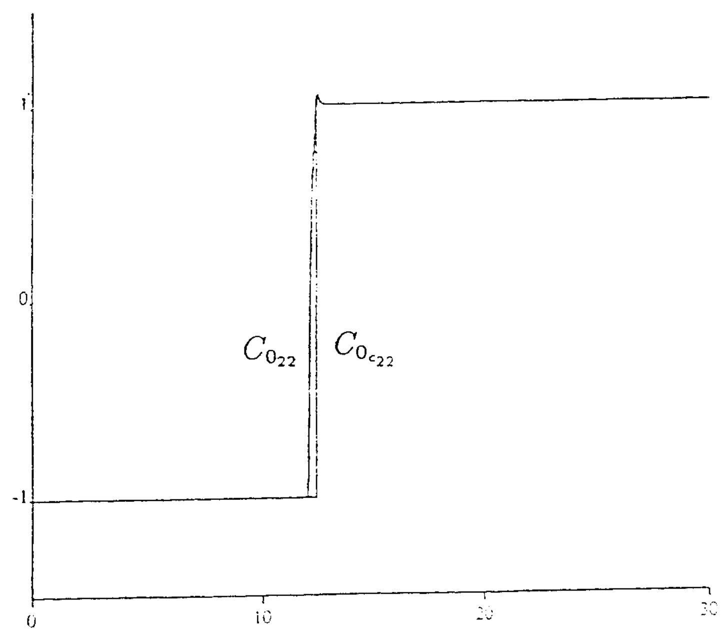

Figure 4.13: $C_{0_{c_{22}}}(t)$ and $C_{0_{22}}(t)$ (with $\epsilon_P = 0.08$)

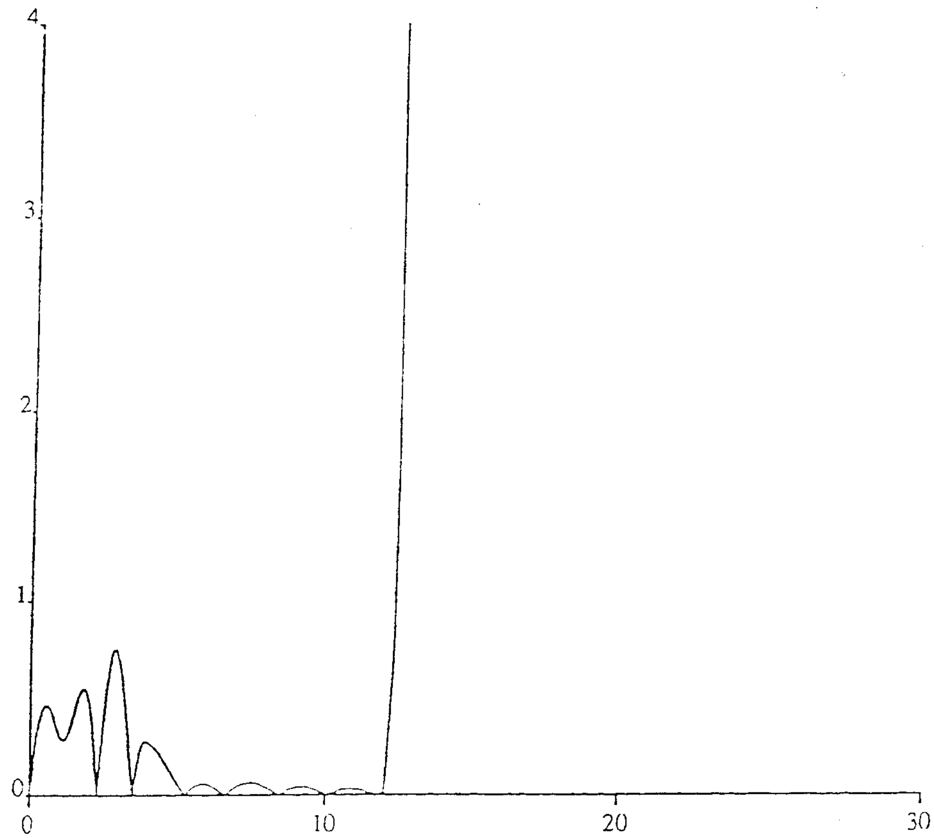

Figure 4.14: $|e_0(t)|$ (with $\epsilon_P = 0$)

Chapter 5

Structural adaptation

5.0 Introduction

In this chapter, we show how the estimation of the structural parameters: Hermite normal form $H(s)$, observability indices $\{\nu_i\}$, and pseudo-observability indices $\{\rho_i\}$, can be introduced in adaptive control.

In chapter 4, we introduced a parameter transformation that removes the requirement of *a priori* knowledge of the high-frequency gain matrix in direct MRAC schemes. However, the *a priori* knowledge of the full interactor matrix or Hermite normal form was still required. Algorithms have appeared in the literature (*e.g.*, Elliot & Wolovich [11], Johansson [13], Dugard *et al.* [12], Dugard *et al.* [28], Scattolini & Clarke [17], Ortega *et al.* [14], Das [7], and Dion *et al.* [50]) reducing this requirement to the *a priori* knowledge of the relative degrees $\{r_i\}$ of the elements on the diagonal of the Hermite normal form (see Definition 1.4). However, when the interactor matrix is not fully known, stability is not proven for the algorithms proposed by Elliot & Wolovich [11], Johansson [13], Dugard *et al.* [28], Scattolini & Clarke [17], Das [7], and Dion *et al.* [50]. The proof of stability proposed in Ortega *et al.* [14] is incorrect. Furthermore, the proof of stability sketched in Dugard *et al.* [12] suffers from serious problems. Indeed, in Dugard *et al.* [12], a modification (identical to the one presented in Goodwin *et al.* [19], Lemma 9.1 and later used in Goodwin & Long [20]) is used in the discrete-time parameter identifier to avoid singularity of the estimate of the time delay gain matrix. This is necessary because its inverse must be computed in the implementation of the control law (similarly to control law (2.3) and (5.2)). However, even if their modification is implemented, the estimate of the time delay gain matrix can still converge arbitrarily close to singularity if the regressor vector is not sufficiently exciting. Therefore, even if it can be proved in theory that the control signal u is bounded, in practice u may not be computable when the estimate of the time delay gain matrix converges close to singularity. Furthermore, this modification cannot be adapted to continuous-time systems. Therefore, for all practical purposes, knowledge of the time delay gain matrix or the high-frequency gain matrix must be assumed.

In this chapter, we present a direct continuous-time MRAC scheme requiring only the knowledge of the relative degrees $\{r_i\}$ of the diagonal elements of the Hermite normal form and an upper bound on the maximum of the observability indices. Global stability is proven without requiring *a priori* knowledge of the high-frequency gain matrix, by using the parameter transformation presented in chapter 4. This result finally makes MIMO MRAC schemes comparable to SISO MRAC schemes in term of applicability and properties.

In chapter 3, it was shown that identifiable parameterizations for direct and indirect adaptive schemes require the knowledge of the observability indices or a set of pseudo-observability indices. If the system is unknown it is likely that this type of knowledge is not available. This has lead researchers to investigate structure estimation techniques. Recent contributions in the domain are by, *e.g.*, Wertz *et al.* [83], Van Overbeek & Ljung [84], Hannan & Kavalieris [85], Eldem & Yildizbayrak [86], Janssen *et al.* [87], and Fuchs [88]. More particularly, Van Overbeek & Ljung [84] have discussed the advantages of using overlapping parameterizations (parameterizations based on the pseudo-observability indices) for structure estimation, assuming that the order of the system is known. They also showed the feasibility of an on-line structure estimation procedure.

In the previous chapter, we extended the results of de Mathelin & Bodson [65] about parameter convergence in recursive identification with canonical structures to recursive identification with pseudo-canonical structures, and found frequency-domain conditions to guarantee the convergence of the parameters, assuming that the selected structure is a valid one (see Theorems 3.4 and 3.5). In this chapter, we reverse the problem and, assuming that sufficient excitation is provided, we find conditions to guarantee that the structure is indeed valid. The results are related to those of Van Overbeek & Ljung [84]. The main difference is that we express minimal conditions on the reference inputs using frequency-domain conditions (white noise inputs are not required). Also, we extend the results to incorporate the case when the order itself is not known. The analysis is carried out in parallel between the canonical and pseudo-canonical forms, allowing for a comparison of their respective advantages. These results could be used for structure selection in an indirect adaptive scheme. Finally, note that these results were also detailed in de Mathelin & Bodson [89].

5.1 Structural adaptation in MRAC

5.1.1 Control algorithm

The following controller structure is considered

$$\begin{aligned} r &= M_0[r_0] \\ u &= KG[r] + \Lambda^{-1}\bar{C}[u] + \Lambda^{-1}\bar{D}[y_p] \end{aligned} \tag{5.1}$$

where $\Lambda(s), \bar{C}(s), \bar{D}(s), K(s) \in \Re^{p\times p}[s]$, $G(s)$ and $M_0(s) \in \Re^{p\times p}(s)$. $M_0(s)$ is a proper stable transfer function matrix, $G(s)$ is a proper stable transfer matrix such that KG is proper, and $\Lambda(s)$ is a Hurwitz diagonal matrix, $\Lambda(s) = \mathrm{diag}\{\lambda_i(s)\}$, such that $\Lambda^{-1}\bar{C}$ and $\Lambda^{-1}\bar{D}$ are proper. Since $\Lambda^{-1}\bar{C}$ is proper, one must still use for implementation the following equivalent control law

$$u = (\bar{C}_0)^{-1}(KG[r] + \Lambda^{-1}(\bar{C} - \Lambda\bar{C}_{0_\infty})[u] + \Lambda^{-1}\bar{D}[y_p]) \tag{5.2}$$

where

$$\bar{C}_{0_\infty} = \lim_{s\to\infty} \Lambda^{-1}\bar{C} \qquad \text{and} \qquad \bar{C}_0 = I - \bar{C}_{0_\infty} \tag{5.3}$$

By combining the equation of the plant, $y_p = P[u]$, with the equation of the controller (5.1), the output y_p can be expressed as

$$y_p = N_R((\Lambda - \bar{C})D_R - \bar{D}N_R)^{-1}\Lambda KG[r] \tag{5.4}$$

where $\{N_R, D_R\}$ is a right MFD of $P(s)$.

Assumptions

(A1) Plant assumptions
The plant is described by a square, nonsingular, strictly proper, and minimum phase transfer function matrix $P(s) \in \Re^{p\times p}(s)$.

(A2) Model assumptions
The reference model $M(s) = G(s)M_0(s)$, where $G(s)$ is a diagonal matrix, $G(s) = \mathrm{diag}\{\frac{1}{(s+a)^{d_i}}\}$ with $a > 0$ and $d_i = (\sum_{k=i}^{p}(r_k - 1)) + 1$, where the integers $\{r_i\}$ are the relative degrees of the diagonal elements of the Hermite normal form of $P(s)$ (see Definition 1.4), and $M_0(s) \in \Re^{p\times p}(s)$ is a proper stable transfer function matrix.

(A3) Reference input assumptions
The reference input, $r_0(t)$, is piecewise continuous and belongs to $L_\infty\ \forall\, t > 0$.

The model output, y_m, is defined as:

$$y_m = GM_0[r_0] = G[r] \tag{5.5}$$

Note that, from assumption (A2), based on the definition of the Hermite normal form and of the interactor matrix (*cf.* Definition 1.4), we have that $H^{-1}G = \xi G$ is a proper, stable, and minimum phase transfer function matrix.

Proposition 5.1 : Extended model reference matching equality
If $\partial\lambda_i = \nu - 1$ and $\nu \geq \nu_{\max}$, $\exists\ \bar{C}^*(s), \bar{D}^*(s), K^* \in \Re^{p\times p}[s]$, solution of the Diophantine equation:

$$N_R[(\Lambda - \bar{C}^*)D_R - \bar{D}^*N_R]^{-1}\Lambda K^*G = G \tag{5.6}$$

such that model matching is achieved (the transfer function from r to y_p is $G(s)$), $\Lambda^{-1}\bar{C}^*$ and $\Lambda^{-1}\bar{D}^*$ are proper. In particular, $I - \bar{C}^*_{0_\infty} = C^*_0 = K_p$ nonsingular, $K^* = H^{-1} = (\Sigma^* F + I)\Delta$, $\partial\bar{D}^* \leq \nu_{\max} - 1$, and $\partial r_i \bar{C}^* \leq \partial\lambda_i$.

Proof

The proof is similar to the proof of Proposition 2.1. The matching equality (5.6) is equivalent to

$$K^* P = I - \Lambda^{-1}\bar{C}^* - \Lambda^{-1}\bar{D}^* P \tag{5.7}$$

Now, let $\{N_L, D_L\}$ be a left coprime MFD of $P(s)$ with D_L row reduced. Using the polynomial matrix division lemma (Lemma 2.2), divide ΛH^{-1} on the right by D_L, then $\exists\ Q(s),\ R(s) \in \Re^{p\times p}[s]$ such that

$$\Lambda H^{-1} = Q D_L + R \qquad \text{and} \qquad R D_L^{-1} \text{ is strictly proper}$$

and $\partial c_i R < \partial c_i D_L \leq \nu_{\max}$. Using Definition 1.4, let

$$\bar{D}^* = -R = Q D_L - \Lambda K_p^{-1} H^{-1} \qquad \bar{C}^* = \Lambda - Q N_L \qquad K^* = H^{-1} = (\Sigma^* F + I)\Delta$$

then, it can be easily verified that the given $\bar{C}^*$, $\bar{D}^*$, and K^* solve the matching equality (5.7). Furthermore, since $\partial\lambda_i = \nu - 1$, $\Lambda^{-1}\bar{D}^*$ is proper. On the other hand,

$$\lim_{s\to\infty} \Lambda^{-1}\bar{C}^* = \bar{C}^*_{0_\infty} = I - \bar{C}^*_0 = \lim_{s\to\infty}(I - H^{-1}P - \Lambda^{-1}\bar{D}^* P) = I - K_p$$

so that $\Lambda^{-1}\bar{C}^*$ is proper and $\partial r_i \bar{C}^* \leq \partial\lambda_i$, *e.g.*, $\partial r_i \bar{C}^* \leq \nu_{\max} - 1$ if $\partial\lambda_i = \nu_{\max} - 1$. □

5.1.2 Adaptation

The matching equality (5.6) is equivalent to (*cf.* Proposition 5.1)

$$I = (\Sigma^* F + I)\Delta P + \Lambda^{-1}\bar{C}^* + \Lambda^{-1}\bar{D}^* P \tag{5.8}$$

Define $L(s) = \mathrm{diag}\{l(s)\}$, with $l(s)$ Hurwitz, $\partial l(s) = d \geq d_1 = (\sum_{k=1}^{p}(r_k - 1)) + 1$ (d_1 is the largest possible degree of all elements of $H^{-1}(s) = \xi(s)$). Then, multiplying both sides of (5.8) by L^{-1} and applying both transfer function matrices to u leads to

$$-L^{-1}\Delta[y_p] = L^{-1}(\Sigma^* F\Delta[y_p] + \Lambda^{-1}(\bar{C}^* - \Lambda)[u] + \Lambda^{-1}\bar{D}^*[y_p]) \tag{5.9}$$

which is an equation where the unknown parameters appear linearly. From Definition 1.4, Σ^* is such that

$$\Sigma^*_{ij}(s) = \begin{cases} 0 & \text{if } i \leq j \\ 0 \text{ or } \sigma_{ij}(s) & \text{if } i > j \end{cases} \qquad \text{with} \quad \partial\sigma_{ij} \leq (\textstyle\sum_{k=j+1}^{i}(r_k - 1)) - 1 \tag{5.10}$$

$$\Rightarrow \quad \partial c_i \Sigma^* \leq \left(\sum_{k=i+1}^{p} (r_k - 1)\right) - 1 = d_i - (r_i + 1) \qquad \forall\, i = 1, \ldots, p-1$$

Therefore, from Proposition 5.1, there exist matrices $\bar{C}_1^*, \ldots, \bar{C}_{\nu-1}^*, \bar{D}_1^*, \ldots, \bar{D}_\nu^* \in \Re^{p\times p}$, matrices $\Sigma_1^*, \ldots, \Sigma_{r_p-1}^* \in \Re^{p\times(p-1)}$, $\Sigma_{r_p}^*, \ldots, \Sigma_{r_p+r_{p-1}-2}^* \in \Re^{p\times(p-2)}, \ldots, \ldots, \Sigma_{d_1-r_1}^* \in \Re^{p\times 1}$, such that

$$\begin{aligned}
(\Lambda L)^{-1}(\bar{C}^* - \Lambda) &= (\bar{C}_{0_\infty}^* - I)\frac{1}{l(s)} + \sum_{i=1}^{\nu-1} \bar{C}_i^* \frac{s^{(i-1)}}{l(s)\lambda(s)} = -\bar{C}_0^* \frac{1}{l(s)} + \sum_{i=1}^{\nu-1} \bar{C}_i^* \frac{s^{(i-1)}}{l(s)\lambda(s)} \\
(\Lambda L)^{-1}\bar{D}^* &= \bar{D}_\nu^* \frac{1}{l(s)} + \sum_{i=1}^{\nu-1} \bar{D}_i^* \frac{s^{(i-1)}}{l(s)\lambda(s)} \\
L^{-1}\Sigma^* F\Delta &= \sum_{i=1}^{d_1-r_1} \Sigma_i^* \frac{s^{(i-1)}(s+a)}{l(s)} T^i \Delta(s)
\end{aligned}$$

where T^i is made of the j first rows of the $p \times p$ identity matrix if $\sum_{k=j+1}^{p}(r_k - 1) \leq i$. Furthermore, given (5.10), several elements of the matrices Σ_i^* will always be identical to zero. For example, if $p = 3$, $r_1 = 1$, $r_2 = 2$, and $r_3 = 3$,

$$\Sigma_1^* = \begin{bmatrix} 0 & 0 \\ x & 0 \\ x & x \end{bmatrix} \quad \text{and} \quad T^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\Sigma_2^* = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ x & x \end{bmatrix} \quad \text{and} \quad T^2 = T^1$$

$$\Sigma_3^* = \begin{bmatrix} 0 \\ 0 \\ x \end{bmatrix} \quad \text{and} \quad T^3 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

where x indicates a possibly nonzero element. The matrix of unknown controller parameters is

$$\theta^{*T} = \begin{bmatrix} \bar{C}_0^* & \ldots & \bar{C}_{\nu-1}^* & \bar{D}_1^* & \ldots & \bar{D}_\nu^* & \Sigma_1^* & \ldots & \Sigma_{d_1-r_1}^* \end{bmatrix}$$

and the regressor vector

$$\begin{aligned}
\psi^T = \Big[& -L^{-1}[u]^T \quad (\Lambda L)^{-1}[u]^T \quad \ldots \quad s^{(\nu-2)}(\Lambda L)^{-1}[u]^T \quad (\Lambda L)^{-1}[y_p]^T \quad \ldots \quad s^{(\nu-2)}(\Lambda L)^{-1}[y_p]^T \\
& L^{-1}[y_p]^T \quad T^1 L^{-1} F\Delta[y_p]^T \quad \ldots \quad s^{(d_1-r_1-1)} T^{d_1-r_1} L^{-1} F\Delta[y_p]^T \Big]
\end{aligned} \tag{5.11}$$

so that (5.9) can be rewritten as

$$-L^{-1}\Delta[y_p] = \theta^{*T}\psi \tag{5.12}$$

Then the following error equation can be derived

$$e_2 = \theta^T\psi + L^{-1}\Delta[y_p] = (\theta^T - \theta^{*T})\psi = \phi^T\psi \tag{5.13}$$

where θ is the estimate of θ^* and ϕ is the parameter error. We will use the normalized least-squares algorithm with covariance resetting of section 2.4 to estimate θ.

The error equation requires the *a priori* knowledge of the relative degrees $\{r_i\}$ of the diagonal elements of the Hermite normal form and an upper bound ν on the observability index $\nu_{\max}$.

5.1.3 Parameter transformation and error formulation

Given the control law (5.2), the matrix of estimated parameters $\bar{C}_0$ must avoid singularity. Therefore, the parameter transformation (4.18) and (4.19) described in Chapter 4 will be implemented. The controller parameters will be denoted θ_c and will be obtained from the estimated parameters θ through the parameter transformation (4.18) and (4.19) (assuming that σ is a lower bound on $|K_p|$ instead of an upper bound as in Chapter 4, since $\bar{C}_0$ is the estimate of K_p instead of of K_p^{-1}). As explained in Chapter 4, $\theta_c(t)$ will be equal to $\theta(t)$ as long as $\bar{C}_0(t)$ is far from singularity.

For the purpose of analyzing stability, we will now represent the adaptive system in terms of its deviation with respect to the ideal situation when $\phi = 0$. Given the definition of the controller parameter matrix θ_c, the control law (5.2) is rewritten as

$$u = \bar{C}_{0_c}^{-1}(KG[r] + \bar{\theta}_c^T \bar{w}) = \bar{C}_{0_c}^{-1}(\Delta G[r] + \tilde{\theta}_c^T w_m^{(3)} + \bar{\theta}_c^T \bar{w}) \tag{5.14}$$

where $\bar{\theta}_c$ and $\tilde{\theta}_c$ are submatrices of θ_c

$$\theta_c^T = \begin{bmatrix} \bar{C}_{0_c} & \bar{\theta}_c^T & \tilde{\theta}_c^T \end{bmatrix} \tag{5.15}$$

and $\bar{w}$ and $w_m^{(3)}$ are defined as

$$\begin{aligned}
\bar{w}^T &= \begin{bmatrix} w^{(1)^T} & w^{(2)^T} & y_p^T \end{bmatrix} \\
w^{(1)^T} &= \begin{bmatrix} \Lambda^{-1}[u]^T & \dots & s^{(\nu-2)}\Lambda^{-1}[u]^T \end{bmatrix} = H_{w^{(1)}u}[u]^T \\
w^{(2)^T} &= \begin{bmatrix} \Lambda^{-1}[y_p]^T & \dots & s^{(\nu-2)}\Lambda^{-1}[y_p]^T \end{bmatrix} = H_{w^{(2)}y_p}[y_p]^T \\
w_m^{(3)^T} &= \begin{bmatrix} T^1 F \Delta G[r]^T & \dots & s^{(d_1-r_1-1)} T^{d_1-r_1} F \Delta G[r]^T \end{bmatrix} = H_{w_m^{(3)} y_m} G[r]^T = H_{w_m^{(3)} y_m}[y_m]^T
\end{aligned}$$

where the transfer function matrices $H_{w^{(1)}u}$, $H_{w^{(2)}y_p}$ are strictly proper and stable and the transfer function matrix $H_{w_m^{(3)} y_m} G$ is proper and stable. From the control law (5.14), the following relationship can be deduced

$$-\Delta[y_m] = -\Delta G[r] = -\bar{C}_{0_c} u + \tilde{\theta}_c^T w_m^{(3)} + \bar{\theta}_c^T \bar{w} = \theta_c^T w \tag{5.16}$$

where

$$w^T = \begin{bmatrix} -u^T & \bar{w}^T & w_m^{(3)^T} \end{bmatrix} \tag{5.17}$$

If we define the signal r_p as

$$r_p = H^{-1}[y_p]$$

then, the matching equality (5.8) applied to u can be rewritten as

$$\bar{C}_0^* u = \Delta[y_p] + \Sigma^* F \Delta[y_p] + \bar{\theta}^{*^T} \bar{w} = r_p + \bar{\theta}^{*^T} \bar{w} \tag{5.18}$$

If we define the *control error*, e_c, as

$$e_c = (\theta_c^T - \theta^{*^T}) w = \phi_c^T w$$

Then, from equations (5.16) and (5.18)

$$e_c = r_p - \Delta G[r] - \bar{\theta}^{*^T} w_m^{(3)} = r_p - (\Delta + \bar{\theta}^{*^T} H_{w_m^{(3)} y_m}) G[r] = r_p - H^{-1} G[r] \tag{5.19}$$

and the *output error*, $e_0 = y_p - y_m$, can be expressed as

$$e_0 = H[r_p] - G[r] = H[\phi_c^T w] = H[e_c] \tag{5.20}$$

Since $y_p = P[u] = H[r_p]$, the control input can also be expressed in terms of the control error as

$$u = P^{-1} H[r_p] = P^{-1} G[r] + P^{-1} H[e_c]$$

Similarly,

$$\bar{w} = \begin{bmatrix} w^{(1)} \\ w^{(2)} \\ y_p \end{bmatrix} = \begin{bmatrix} H_{w^{(1)} u} P^{-1} H \\ H_{w^{(2)} y_p} H \\ H \end{bmatrix} [r_p] = H_{\bar{w} r_p}[r_p] \tag{5.21}$$

$$\psi = \begin{bmatrix} -l^{-1} P^{-1} \\ l^{-1} H_{w^{(1)} u} P^{-1} \\ l^{-1} H_{w^{(2)} y_p} \\ L^{-1} \\ l^{-1} H_{w_m^{(3)} y_m} \end{bmatrix} [y_p] = H_{\psi y_p}[y_p] = H_{\psi y_p} H[r_p] = H_{\psi r_p}[r_p] \tag{5.22}$$

where the transfer function matrices $H_{\bar{w} r_p}$, $H_{\psi r_p}$, are stable and strictly proper and stable, and $H_{\psi y_p}$ is stable and proper. If we define the model signals, ψ_m, as the signals ψ when $\phi = 0$

$$\psi_m = H_{\psi y_p}[y_m] = H_{\psi y_p} G[r] = H_{\psi_m r}[r] \tag{5.23}$$

then, we can define the *regressor error*, e_ψ, as

$$e_\psi = \psi - \psi_m = H_{\psi y_p}[e_0] = H_{\psi r_p}[e_c] \tag{5.24}$$

5.1.4 Stability

The following theorem is the main stability theorem. It shows that given any initial condition and any bounded input $r(t)$, the states of the adaptive system remain bounded and the output error converges to zero as $t \to \infty$. Furthermore, the difference between the regressor vector ψ and the corresponding model vector ψ_m converges to zero as $t \to \infty$.

Theorem 5.1 : Global stability of the extended direct MRAC system - K_p unknown

Consider the MIMO MRAC system described in Section 5.1.2 Assume that the parameter estimation algorithm (2.29) and the hysteresis transformation (4.18) and (4.19) are used. If the reference input $r \in L_\infty$ and is piecewise continuous, then

- All states of the adaptive system are bounded functions of time.
- The output error $e_0 = y_p - y_m \in L_\infty$ and $\lim_{t\to\infty} e_0 = 0$.
- The regressor error $e_\psi = \psi - \psi_m \in L_\infty$ and $\lim_{t\to\infty} e_\psi = 0$.

For the stability proof, we follow similar steps as for the proof of Theorem 4.1.

Proof of Theorem 5.1

A. All signals belong to $L_{\infty e}$
Similarly, to section 4.3.2, it can be shown that $(y_p, w^{(1)}, w^{(2)}, \psi)$ belong to $L_{\infty e}$ and, therefore, $(\bar{w},\ w,\ u,\ r_p)$ belong to $L_{\infty e}$.

B. ψ is regular and all signals are bounded by $\|\psi_t\|_\infty$
Let $x, \dot{x}, \Re_+ \to \Re^n \in L_{\infty e}$. Then, we say x is *regular* iff $\exists k_1, k_2 > 0$ such that

$$|\dot{x}(t)| \leq k_1 \|x_t\|_\infty + k_2$$

Recall from (5.19) and (5.14) that

$$\begin{aligned} r_p &= \bar{C}_0^* u - \bar{\theta}_c^{*T} \bar{w} \\ &= \bar{C}_0^* \bar{C}_{0_c}^{-1} K G[r] + (\bar{C}_0^* \bar{C}_{0_c}^{-1} \bar{\theta}_c^T - \bar{\theta}_c^{*T}) \bar{w} \end{aligned} \tag{5.25}$$

and from (5.21), that

$$\bar{w} = H_{\bar{w}r_p}[r_p]$$

with $H_{\bar{w}r_p}$ stable and strictly proper. Therefore, applying Lemma 4.4

$$\begin{aligned} |\bar{w}| &\leq k\|r_{p_t}\|_\infty + |\epsilon(t)| \leq \|\bar{w}_t\|_\infty + k \\ |\frac{d}{dt}\bar{w}| &\leq k\|\bar{w}_t\|_\infty + k \end{aligned}$$

for some constant $k > 0$ and where ϵ is an exponentially decreasing term. Again we use the single symbol k, whenever an inequality is valid for some positive constant. Therefore, $\bar{w}$ is regular. Now, recall from (4.27) that

$$\psi = H_{\psi r_p}[r_p]$$

with $H_{\psi r_p}$ strictly proper and stable. Therefore, by applying Lemma 4.4

$$\begin{aligned} |\psi| &\leq k\|r_{p_t}\|_\infty + |\epsilon(t)| \leq k\|\bar{w}_t\|_\infty + k \\ |\psi| &\leq k\|F^{-1}[r_p]_t\|_\infty + |\epsilon(t)| \\ |\frac{d}{dt}\psi| &\leq k\|r_{p_t}\|_\infty + |\epsilon(t)| \leq k\|\bar{w}_t\|_\infty + k \end{aligned} \tag{5.26}$$

Now, define $\bar{\psi}$ as $l^{-1}[\bar{w}]$. Since $\bar{w}$ is regular, l^{-1} is strictly proper and minimum phase, by applying Lemma 4.5

$$|\bar{w}| \leq k\|\bar{\psi}_t\|_\infty + k$$

Since

$$\psi = \begin{bmatrix} -L^{-1}P^{-1}H[r_p] \\ \bar{\psi} \\ l^{-1}H_{w_m^{(3)}y_m}H[r_p] \end{bmatrix}$$

$|\bar{\psi}| \leq |\psi|$ and

$$\begin{aligned} |\bar{w}| &\leq k\|\psi_t\|_\infty + k \\ |\frac{d}{dt}\psi| &\leq k\|\psi_t\|_\infty + k \end{aligned}$$

so that ψ is regular. Now, from (5.25)

$$|r_p| \leq k|\bar{w}| + k \leq k\|\psi_t\|_\infty + k$$

Similarly, from (5.14)

$$|u| \leq k|\bar{w}| + k \leq k\|\psi_t\|_\infty + k$$

and applying Lemma 4.4 to $y_p = P[u]$

$$|y_p| \leq k\|\psi_t\|_\infty + k$$

Consequently,

$$|u|, |y_p|, |r_p|, |\bar{w}|, |w|, \leq k\|\psi_t\|_\infty + k$$

and if ψ is proved to be bounded, then all signals will be bounded.

C. $\underline{\lim_{t\to\infty} \beta_c = 0}$

This comes from the regularity of ψ. The proof is identical to Part C in the proof of Theorem 4.1.

D. Relationship between output error and identification error

Recall from equation (5.20) that

$$e_0 = y_p - y_m = H[\phi_c^T w]$$

therefore, from (5.15), (5.17), and (5.22)

$$\begin{aligned}
(HL)^{-1}[e_0] &= L^{-1}(\Sigma^* F + I)\Delta)[e_0] = L^{-1}[\phi_c^T w] \\
&= \phi_c^T \psi + (L^{-1}[\phi_c^T w] - \phi_c^T \psi) \\
&= \phi_c^T \psi + (L^{-1}[\phi_c^T w] - \phi_c^T l^{-1}[w]) + (\tilde{\theta}_c - \tilde{\theta}^*)^T l^{-1} H_{w_m^{(3)} y_m}[y_m - y_p]
\end{aligned}$$

which can be rewritten

$$(\tilde{\theta}_c^T H_{w_m^{(3)} y_m} + \Delta) L^{-1}[e_0] = \phi_c^T \psi + (L^{-1}[\phi_c^T w] - \phi_c^T l^{-1}[w]) \tag{5.27}$$

Now, this can be used to transfer convergence properties from the identifier error to the output error.

E. All signals are bounded, $\lim_{t\to\infty} e_0 = 0$, and $\lim_{t\to\infty} e_\psi = 0$

Suppose that $\{A, B, C\}$ is a minimal realization of $l^{-1}(s)$, *i.e.*, $C(sI - A)^{-1}B = l^{-1}(s)$. Let x and W such that

$$\begin{aligned}
\dot{x} &= Ax + Bw^T \phi_c \\
\dot{W} &= AW + Bw^T
\end{aligned}$$

then

$$\begin{aligned}
Cx &= l^{-1}[w^T \phi_c] \\
CW &= l^{-1}[w^T]
\end{aligned}$$

so that

$$L^{-1}[\phi_c^T w] - \phi_c^T l^{-1}[w] = (C(x - W\phi_c))^T \tag{5.28}$$

Furthermore, $\forall t \notin \mathcal{T} = \{t_k\} \cup \{T_k\} \cup \{\tau_k\} \cup \{t_{ij}^k\}$

$$\frac{d}{dt}(x - W\phi_c) = A(x - W\phi_c) - W\frac{d}{dt}(\phi_c) \tag{5.29}$$

where $\forall t \notin \mathcal{T}$

$$\frac{d}{dt}\phi_c = \dot{\phi} + \dot{P}Q + P\dot{Q} \tag{5.30}$$

and $\forall t \notin \mathcal{T}$

$$\dot{Q} = \begin{cases} 0 & \text{if } T_k < t < t_{k+1} \\ -g\frac{P\psi\psi^T Q}{1+\gamma\psi^T\psi} + gP_p(P_p^T P_p)^{-1}P_p^T\frac{[P\psi\psi^T+\psi\psi^T P]}{1+\gamma\psi^T\psi}Q - P_p(P_p^T P_p)^{-1}\dot{\bar{C}}_0^T & \text{if } \begin{cases} t_k < t < T_k \; \forall\, k \\ \sigma_{\min}(P_p^T(t)) > \epsilon_P \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

Define ζ such that

$$\zeta = \begin{cases} \dot{Q} & \text{if } \begin{cases} t_k < t < T_k \;\; \forall k \\ \sigma_{\min}(P_p^T(t)) > \epsilon_P \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

Then, given the properties of the identifier (*cf.* Lemma 2.3) when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, given the fact that $\exists\, T > 0$ such that $\bar{C}_{0_c}(t) = \bar{C}_0(t)$ and $\zeta(t) = 0 \;\forall\, t > T$ when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, and given the fact that $(P_p^T P_p)^{-1}$ is bounded if $\sigma_{\min}(P_p^T) > \epsilon_P$, $\zeta \in L_2 \cap L_\infty$. Given the definition of ζ, from (5.30)

$$\frac{d}{dt}\phi_c = \dot{\phi} + \dot{P}Q + \zeta \qquad \forall t \notin \mathcal{T} \tag{5.31}$$

Define Δ_{T_k}, Δ_{τ_k}, and $\Delta_{t_{ij}^k}$, as

$$\begin{aligned} \Delta_{T_k} &= (\phi_c)_{|t=T_k} - (\phi_c)_{|t=T_k^-} \\ \Delta_{\tau_k} &= (\phi_c)_{|t=\tau_k} - (\phi_c)_{|t=\tau_k^-} \\ \Delta_{t_{ij}^k} &= (\phi_c)_{|t=t_{ij}^k} - (\phi_c)_{|t=t_{ij}^{k-}} \end{aligned} \tag{5.32}$$

Finally, combining together (5.27), (5.28), (5.29), (5.31), and (5.32)

$$\begin{aligned} (\tilde{\theta}_c^T H_{w_m^{(3)} y_m} + \Delta)L^{-1}[e_0] &= \phi_c^T\psi + (Ce^{At}(x - W\phi_c)_{|t=0})^T \\ &\quad -\left(\int_0^t Ce^{A(t-\tau)}W(\tau)(\dot{\phi}(\tau) + \dot{P}(\tau)Q(\tau) + \zeta(\tau))d\tau\right)^T \\ &\quad -\Big(\sum_{T_k<t} Ce^{A(t-T_k)}W(T_k)\Delta_{T_k}s(t-T_k)\Big)^T \\ &\quad -\Big(\sum_{\tau_k<t} Ce^{A(t-\tau_k)}W(\tau_k)\Delta_{\tau_k}s(t-\tau_k)\Big)^T \\ &\quad -\Big(\sum_{t_{ij}^k<t} Ce^{A(t-t_{ij}^k)}W(t_{ij}^k)\Delta_{t_{ij}^k}s(t-t_{ij}^k)\Big)^T \end{aligned} \tag{5.33}$$

where $s(t)$ is the unit step function. By Lemma 4.4, $|W| \le k\|w_t\|_\infty + |\epsilon(t)| \le k\|\psi_t\|_\infty + k$. Since $(\dot{\phi} + \dot{P}Q + \zeta) \in L_2 \cap L_\infty$, Lemma 4.6 can be applied. Further, $\{T_k\}$ and $\{t_{0j}^k\}$ are finite

sets, and $t_{ij}^k = t_{0j}^k + i\delta_t$, $\forall\, i$. Also, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, $\exists\, k^* \geq 0$ such that $\Delta_{\tau_k} = 0$, $\forall\, k > k^*$, and $\{t_{ij}^k\}$ is a finite set. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, then $\{\tau_k\}$ is a finite set, $\lim_{t\to\infty}\phi(t) = \phi_\infty$, $\lim_{t\to\infty} P(t) = P_\infty$, and, from Lemma 4.2, $\lim_{i\to\infty} |\Delta_{t_{ij}^k}| = 0$. Therefore, similarly to Part E in the proof of Theorem 4.1, (5.33) may be put under the following form

$$(\tilde{\theta}_c^T H_{w_m^{(3)} y_m} + \Delta)L^{-1}[e_0] = \tilde{\beta}(1 + \|\psi_t\|_\infty) \tag{5.34}$$

with $\lim_{t\to\infty}\tilde{\beta} = 0$. From (5.25) and (5.26)

$$|\frac{d}{dt}F^{-1}[r_p]| \leq k\|\bar{w}_t\|_\infty + k \leq k\|\psi_t\|_\infty + k \leq k\|F^{-1}[r_p]_t\|_\infty + k \tag{5.35}$$

so that $F^{-1}[r_p]$ is regular. Since $y_p = FHF^{-1}[r_p]$, with FH stable, proper, and minimum phase, applying Lemma 4.5

$$|F^{-1}[r_p]| \leq k\|y_{p_t}\|_\infty + k \leq k\|e_{0_t}\|_\infty + k \tag{5.36}$$

By applying Lemma 4.4 to (5.20)

$$|\frac{d}{dt}e_0| \leq k\|w_t\|_\infty + |\epsilon(t)|$$

Then, from (5.35) and (5.36)

$$|\frac{d}{dt}e_0| \leq k\|F^{-1}[r_p]_t\|_\infty + k \leq k\|e_{0_t}\|_\infty + k$$

so that e_0 is regular. Therefore, since L^{-1} is stable, strictly proper, and minimum phase, using Lemma 4.5

$$|e_0| \leq k\|L^{-1}[e_0]_t\|_\infty + k \tag{5.37}$$

Let $v = (\tilde{\theta}_c^T H_{w_m^{(3)} y_m} + \Delta)L^{-1}[e_0]$. Then, from Sections 5.1.1 and 5.1.2, v has the following structure

$$\begin{aligned}
v_1 &= \frac{(s+a)^{r_1}}{l(s)}[e_{0_1}] \\
v_2 &= \sum_{i=0}^{r_2-2} \tilde{\theta}_{c_{2j_{21}(i)}}^T \frac{s^i(s+a)^{r_1+1}}{l(s)}[e_{0_1}] + \frac{(s+a)^{r_2}}{l(s)}[e_{0_2}] \\
v_3 &= \sum_{i=0}^{r_2+r_3-3} \tilde{\theta}_{c_{3j_{31}(i)}}^T \frac{s^i(s+a)^{r_1+1}}{l(s)}[e_{0_1}] + \sum_{i=0}^{r_3-2} \tilde{\theta}_{c_{3j_{32}(i)}}^T \frac{s^i(s+a)^{r_2+1}}{l(s)}[e_{0_2}] + \frac{(s+a)^{r_3}}{l(s)}[e_{0_3}] \\
&\vdots \\
v_p &= \sum_{i=0}^{d_1-(r_1+1)} \tilde{\theta}_{c_{pj_{p1}(i)}}^T \frac{s^i(s+a)^{r_1+1}}{l(s)}[e_{0_1}] + \sum_{i=0}^{d_2-(r_2+1)} \tilde{\theta}_{c_{pj_{p2}(i)}}^T \frac{s^i(s+a)^{r_2+1}}{l(s)}[e_{0_2}] + \ldots + \frac{(s+a)^{r_p}}{l(s)}[e_{0_p}]
\end{aligned} \tag{5.38}$$

where $j_{kl}(i)$ is the position in the k-th row of $\tilde{\theta}_c^T$ of the estimate of the coefficient of i-th power in $\sigma_{kl}(s)$. Then, after applying the linear operator $\frac{1}{(s+a)^{r_i}}$ to the i-th element of v for $i = 1, \ldots, p$, using Lemma 4.4 and (5.37) yield

$$\begin{aligned}
|l^{-1}[e_{0_1}]| &\leq k\|v_{1_t}\|_\infty + |\epsilon(t)| \\
|l^{-1}[e_{0_2}]| &\leq k\|e_{0_{1_t}}\|_\infty + k\|v_{2_t}\|_\infty + |\epsilon(t)| \leq k\|v_{1_t}\|_\infty + k\|v_{2_t}\|_\infty + k \\
|l^{-1}[e_{0_3}]| &\leq k\|e_{0_{1_t}}\|_\infty + k\|e_{0_{2_t}}\|_\infty + k\|v_{3_t}\|_\infty + |\epsilon(t)| \leq k\|v_{1_t}\|_\infty + k\|v_{2_t}\|_\infty + k\|v_{3_t}\|_\infty + k \\
&\vdots \\
|l^{-1}[e_{0_p}]| &\leq k\|e_{0_{1_t}}\|_\infty + \ldots + k\|e_{0_{(p-1)_t}}\|_\infty + k\|v_{p_t}\|_\infty + |\epsilon(t)| \leq k\|v_{1_t}\|_\infty + \ldots + k\|v_{p_t}\|_\infty + k
\end{aligned}$$

Therefore,

$$|L^{-1}[e_0]| \leq k\|v_t\|_\infty + k \tag{5.39}$$

Combining (5.26), (5.36), and (5.37) yields

$$|\psi| \leq k\|r_{p_t}\|_\infty + |\epsilon(t)| \leq k\|e_{0_t}\|_\infty + k \leq k\|L^{-1}[e_0]_t\|_\infty + k \tag{5.40}$$

Then, from (5.34), (5.39) and (5.40), it can be deduced that

$$|v| \leq |\tilde{\beta}|(1 + \|\psi_t\|_\infty) \leq k|\tilde{\beta}|(1 + \|v_t\|_\infty) \tag{5.41}$$

In this expression, $\lim_{t\to\infty} \tilde{\beta} = 0$. Consequently, $\exists T > 0$ such that $\forall t > T$, $k|\tilde{\beta}(t)| < 1$. Then

$$|v(t)| \leq k\|v_T\|_\infty + k \qquad \forall\, t > T$$

This can be interpreted as an application of the small gain theorem (*cf.* Sastry & Bodson [3], pp. 149). Since $v \in L_{\infty e}$, it follows that $v \in L_\infty$. Then, from (5.41)

$$|v| \leq k|\tilde{\beta}|$$

Therefore, $\lim_{t\to\infty} v = 0$. From (5.39), it follows that $L^{-1}[e_0] \in L_\infty$. Therefore, from (5.40), $\psi \in L_\infty$ and all signals in the adaptive system are bounded. Since, e_0 is regular, $\dot{e}_0 \in L_\infty$. Given the value of the first element of v (*cf.* (5.38)), we have that $\frac{(s+a)^{r_1}}{l(s)}[e_{0_1}] \in L_\infty$ and $\lim_{t\to\infty} \frac{(s+a)^{r_1}}{l(s)}[e_{0_1}] = 0$. Then, from Lemma 4.7, $\lim_{t\to\infty} e_{0_1} = 0$. Given the value of the second element of v (*cf.* (5.38)), applying Lemma 4.6 we have that $\frac{(s+a)^{r_2}}{l(s)}[e_{0_2}] \in L_\infty$ and $\lim_{t\to\infty} \frac{(s+a)^{r_2}}{l(s)}[e_{0_2}] = 0$. Then, from Lemma 4.7, $\lim_{t\to\infty} e_{0_2} = 0$. Similarly, it can be shown that $\lim_{t\to\infty} e_{0_i} = 0$, $i = 3, \ldots, p$, so that $\lim_{t\to\infty} e_0 = 0$. Finally, since $e_\psi = H_{\psi y_p}[e_0]$ (*cf.* (5.24)) with $H_{\psi y_p}$ stable and proper, by Lemma 4.6, we have that $e_\psi \in L_\infty$ and $\lim_{t\to\infty} e_\psi = 0$. □

Finally, note that as in Theorem 4.1 there are no requirements on the reference signals r except, of course, that they must be bounded and piecewise continuous.

5.2 Structure selection in recursive identification

In section 3.2.3, we assumed that the observability indices or a set of pseudo-observability indices were known to derive a recursive identifier for an identifiable parameterization. If this knowledge is not available, we would like to know how to assess if a correct or incorrect structure has been selected. In this section, we show how the autocorrelation of the regressor vector ψ can be used as a criteria for structure selection. The main results reside in the following two theorems.

Theorem 5.2 : Structure selection - Identification of pseudo-canonical left MFD
Given a strictly proper stable multivariable system with unknown pseudo-observability indices. Assume that the inputs are sufficiently rich for identification of the system (for example, assume that the sufficient condition of Theorem 3.5 is respected).

1. Let the true system be of order n, the structure defined by the pseudo-observability indices $\{\rho_i\}$ with $\sum_{i=1}^{p} \rho_i = n$ is a valid structure if and only if

$$R_\psi(0) > 0$$

   where the regressor ψ is given by (3.11).

2. The structure defined by the pseudo-observability indices $\{\rho_i\}$ is a valid structure if and only if

$$R_\psi(0) > 0 \quad \text{and} \quad R_{w_i}(0) = 0 \qquad \forall\, i = 1, \ldots, p$$

   where ψ is given by (3.11) and w_i is the augmented vector

$$w_i = \begin{bmatrix} \psi \\ \frac{y_{p_i}}{(s+a)^{\rho_{\max}-\rho_i}} \end{bmatrix}$$

Proof

Part 1:
First, let us show that if the set of indices $\{\rho_i\}$ defines a valid structure of order n then $R_\psi(0) > 0$. Suppose it was not true. Then there would exist a vector $k \in \Re^{n+m\rho_{\max}}$, $k \neq 0$ such that $k^T R_\psi(0) = [0]$. Since the input is sufficiently rich (by having, for example, $n + \rho_{\max}$ distinct frequency components per input), this is equivalent to:

$$k^T H_{\psi u}(s) = [0] \quad \forall s \tag{5.42}$$

where $H_{\psi u}(s)$ is the transfer function between the input u and the regressor vector ψ

$$H_{\psi u}(s) = \begin{bmatrix} \frac{1}{(s+a)} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{1}{(s+a)^{\rho_{\max}}} & 0 & & 0 \\ 0 & \frac{1}{(s+a)} & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \frac{1}{(s+a)^{\rho_{\max}}} & & 0 \\ & & \ddots & \\ 0 & 0 & & \frac{1}{(s+a)} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \frac{1}{(s+a)^{\rho_{\max}}} \\ \frac{P_{11}(s)}{(s+a)^{\rho_{\max}-\rho_1+1}} & \frac{P_{12}(s)}{(s+a)^{\rho_{\max}-\rho_1+1}} & \cdots & \frac{P_{1m}(s)}{(s+a)^{\rho_{\max}-\rho_1+1}} \\ \vdots & \vdots & & \vdots \\ \frac{P_{11}(s)}{(s+a)^{\rho_{\max}}} & \frac{P_{12}(s)}{(s+a)^{\rho_{\max}}} & \cdots & \frac{P_{1m}(s)}{(s+a)^{\rho_{\max}}} \\ \frac{P_{21}(s)}{(s+a)^{\rho_{\max}-\rho_2+1}} & \frac{P_{22}(s)}{(s+a)^{\rho_{\max}-\rho_2+1}} & \cdots & \frac{P_{2m}(s)}{(s+a)^{\rho_{\max}-\rho_2+1}} \\ \vdots & \vdots & & \vdots \\ \frac{P_{p1}(s)}{(s+a)^{\rho_{\max}}} & \frac{P_{p2}(s)}{(s+a)^{\rho_{\max}}} & \cdots & \frac{P_{pm}(s)}{(s+a)^{\rho_{\max}}} \end{bmatrix} \tag{5.43}$$

Therefore, $\exists\, N(s) \in \Re^{1\times m}[s]$, $D(s) \in \Re^{1\times p}[s]$, $\neq 0$, such that (5.42) can be rewritten as

$$N(s) - D(s)P(s) = [0] \quad \text{with} \quad \begin{cases} \partial c_i D < \rho_i \\ \partial N < \rho_{\max} \end{cases}$$

Define $\{N_L(s), D_L(s)\}$ as the pseudo-canonical left MFD with pseudo-observability indices $\{\rho_i\}$. Then $\{N_L(s), D_L(s)\}$ is unique and $N_L(s) - D_L(s)P(s) = 0$. Now, define $\{\bar{N}_L(s), \bar{D}_L(s)\}$ such that

$$\begin{aligned} [\bar{N}_L(s) \quad \bar{D}_L(s)]_i &= [N_L(s) \quad D_L(s)]_i \quad \text{if} \quad \rho_i \neq \rho_{\max} \\ &= [N_L(s) \quad D_L(s)]_i + [N(s) \quad D(s)] \quad \text{if} \quad \rho_i = \rho_{\max} \end{aligned}$$

where $[\ \ ]_i$ is the i-th row of the matrix $[\ \ ]$. Therefore, $\{\bar{N}_L(s), \bar{D}_L(s)\}$ is another pseudo-canonical form with pseudo-observability indices $\{\rho_i\}$, which is impossible.

The next step consists in showing that if $R_{\psi u}(0) > 0$ then the set of indices $\{\rho_i\}$ defines a valid structure or, equivalently, if the $\{\rho_i\}$ are not a valid set of pseudo-observability indices then $R_{\psi u}(0) = 0$. Suppose that $\{\rho_i\}$ is not a valid set of pseudo-observability indices then there must exist an index l, $1 \leq l \leq p$ such that, for all state-space realizations $\{A, B, C\}$ of the

system, the matrix $K[C,A]$

$$K[C,A] = \begin{bmatrix} C_1 \\ \vdots \\ C_1 A^{r_1-1} \\ C_2 \\ \vdots \\ C_p A^{r_p-1} \end{bmatrix} \qquad \text{with} \qquad r_i = \begin{cases} \min(\rho_i, \bar{\rho}_l + 1) & \text{if } i < l \\ \min(\rho_i, \bar{\rho}_l) & \text{if } i \geq l \\ \bar{\rho}_l \leq \rho_l - 1 & \end{cases} \tag{5.44}$$

has full rank $r = \sum_{i=1}^{p} r_i$, but such that the row vector $C_l A^{\bar{\rho}_l}$ is linearly dependent on the rows of the matrix $K[C,A]$, *i.e.*,

$$C_l A^{\bar{\rho}_l} = \sum_{j=1}^{p} \sum_{k=0}^{r_j-1} \alpha_{ljk} C_j A^k$$

Since $K[C,A]$ has full rank r, there must exist a set of pseudo-observability indices $\{\bar{\rho}_i\}$ such that

$$\begin{aligned} \bar{\rho}_i &\geq \min(\rho_i, \bar{\rho}_l + 1) \quad \text{if} \quad i < l \\ \bar{\rho}_i &\geq \min(\rho_i, \bar{\rho}_l) \quad \text{if} \quad i > l \\ \bar{\rho}_l &\leq \rho_l - 1 \end{aligned}$$

Suppose that $\{\bar{N}_L(s), \bar{D}_L(s)\}$ is the pseudo-canonical left MFD corresponding to the indices $\{\bar{\rho}_i\}$. Then, the l-th row of $\bar{D}_L(s)$ is such that (*cf.* Definition 1.12)

$$\partial \bar{D}_{L_{li}} \begin{cases} < r_i & \text{if} \quad i \neq l \\ = \bar{\rho}_l & \text{if} \quad i = l \end{cases} \qquad \Rightarrow \qquad \partial \bar{D}_{L_{li}} < \rho_i \quad \forall i$$

and, since

$$\bar{N}_{L_l}(s) - \bar{D}_{L_l}(s) P(s) = 0$$

given the definition of $H_{\psi u}(s)$ (5.43),

$$\exists \;\; k \in \Re^{n + m\rho_{max}}, \; k \neq 0 \quad \text{such that} \quad k^T H_{\psi u}(s) = [0] \;\; \forall s \qquad \Rightarrow \qquad R_{\psi u}(0) = 0$$

Part 2:
Using the same technique than in Part 1, it can be shown that if the $\{\rho_i\}$ form a valid set of pseudo-observability indices then $R_{\psi u}(0) > 0$ if the inputs are sufficiently rich. Furthermore, since $\tilde{y}_p = \theta^{*T}\psi$, there exist p vectors $k_i \neq 0$ such that $k_i^T w_i = 0$, so that $R_{w_i}(0) = 0 \; \forall i = 1, \ldots, p$.

As in Part 1, the next step consist in proving that if the $\{\rho_i\}$ are not a valid set of pseudo-observability indices then $R_{\psi u}(0) = 0$ or $\exists \; i$ such that $R_{w_i u}(0) > 0$. If $\{\rho_i\}$ is not a valid

set of pseudo-observability indices then only two different situations may arise. In the first situation, there exists an index l, $1 \leq l \leq p$ such that, for all state-space realizations $\{A, B, C\}$ of the system, the matrix $K[C, A]$ as defined in Part 1 (see equation (5.44)), has full rank $r = \sum_{i=1}^{p} r_i$, but the row vector $C_l A^{\bar{\rho}_l}$ is linearly dependent on the rows of the matrix $K[C, A]$. Then, similarly to Part 1, it can be shown that $R_{\psi u}(0) = 0$. In the second situation, there is a valid set of pseudo-observability indices $\{\bar{\rho}_i\}$ and an index l, $1 \leq l \leq p$, such that

$$\bar{\rho}_i \begin{cases} > \rho_l & \text{if} \quad i = l \\ \geq \rho_i & \text{if} \quad i \neq l \end{cases}$$

Then, it can be shown that $R_{w_l} > 0$. Indeed, suppose it was not true. There would exist a vector $k \neq 0$ such that $k^T R_{w_l}(0) = [0]$ or, since the input is assumed to be sufficiently rich, such that $k^T H_{w_l u}(s) = [0]$, $\forall s$, where $H_{w_l u}(s)$ is given by

$$H_{w_l u}(s) = \begin{bmatrix} & H_{\psi u}(s) & \\ \frac{P_{l1}(s)}{(s+a)^{\rho_{\max}-\rho_l}} & \cdots & \frac{P_{lm}(s)}{(s+a)^{\rho_{\max}-\rho_l}} \end{bmatrix}$$

In other words, there would exist $N(s) \in \Re^{1\times m}[s]$, $D(s) \in \Re^{1\times p}[s]$ such that

$$N(s) - D(s)P(s) = [0] \qquad \text{and} \qquad \begin{cases} \partial c_i D < \rho_i & i \neq l \\ \partial c_l D \leq \rho_l \\ \partial N < \rho_{\max} \end{cases}$$

Let $\{N_L(s), D_L(s)\}$ be the pseudo-canonical left MFD with indices $\{\bar{\rho}_i\}$. Now define $\{\bar{N}_L(s), \bar{D}_L(s)\}$ such that

$$\begin{aligned} [\bar{N}_L(s) \quad \bar{D}_L(s)]_i &= [N_L(s) \quad D_L(s)]_i \quad \text{if} \quad i \neq l \\ &= [N_L(s) \quad D_L(s)]_i + [N(s) \quad D(s)] \quad \text{if} \quad i = l \end{aligned}$$

Then $\{\bar{N}_L(s), \bar{D}_L(s)\}$ is another pseudo-canonical form with pseudo-observability indices $\{\bar{\rho}_i\}$, which is impossible. □

Theorem 5.3 : Structure selection - Identification of canonical left MFD

Given a strictly proper stable multivariable system with unknown observability indices. Assume that the inputs are sufficiently rich for identification of the system (for example, assume that the sufficient condition of Theorem 3.4 is respected).

The structure defined by the observability indices $\{\nu_i\}$ is the correct structure if and only if

$$R_{\psi_i}(0) > 0 \quad \text{and} \quad R_{w_i}(0) = 0 \qquad \forall\, i = 1, \ldots, p$$

where ψ_i is given by (3.15) and w_i is the augmented vector

$$w_i = \begin{bmatrix} \psi_i \\ y_{p_i} \end{bmatrix}$$

The proof of Theorem 5.3 can be easily deduced from the proof of Theorem 5.2. By properly modifying the proof of Theorem 5.2, Part 2, it can be shown that, assuming sufficiently rich inputs, the indices $\{\nu_i\}$ are the observability indices of the system if and only if $R_{\psi_i u}(0) > 0$ and $R_{w_i}(0) = 0 \; \forall i = 1, \ldots, p$.

Remarks:

In Section 3.3.3, we had found conditions to guarantee persistency of excitation of the regressor, *assuming that the correct structure is known.* Here, we have reversed the question, and asked for conditions to validate the structure, assuming that the input is sufficiently rich to guarantee persistency of excitation for the correct structure.

If the order of the system is known, the pseudo-canonical structure has some advantages over the canonical structure. Indeed, $R_\psi(0)$ alone can be used to test that a pseudo-canonical structure has been chosen. Not only is $R_\psi(0) > 0$ for the correct structure, but it is also *not* positive definite for incorrect structure, no matter how rich the reference input is. Another advantage of the pseudo-canonical parameterizations is the fact that they are overlapping. In the sense that two pseudo-canonical parameterizations of same order define two overlapping classes of systems. Consequently, it is possible to change structure in a recursive identification algorithm without losing the information accumulated in the parameter estimates (by applying to a state-space representation of the estimated system a similarity transformation from the old to the new structure). This property of pseudo-canonical structure was used by Van Overbeek & Ljung [84] to design an on-line structure selection algorithm for multivariable stochastic systems of known order and without measurable inputs.

The disadvantage of pseudo-canonical forms is that more parameters need to be identified and complex constraints must be added in the identification algorithm to guarantee strict properness of the estimate.

Finally, it should be pointed out that if a least-squares algorithm without covariance resetting ($k_1 = 0$) and without normalization ($\gamma = 0$) is used for identification (*cf.* Section 2.4), using averaging theory, it can be shown that the covariance matrix P of the least-squares algorithm converges to a neighborhood of $(gR_\psi(0))^{-1}$ if g is sufficiently small, so that P can be used to check the correctness of the identifier structure.

Chapter 6

Conclusions

6.1 Contributions

This thesis presented new analytical results on parameterization, stability, and convergence of multivariable adaptive control systems. It also showed new algorithms to relax requirements on *a priori* information.

The thesis focused on direct adaptive control algorithms and recursive identification algorithms for deterministic multivariable systems. A first contribution was the modification of current algorithms in order to ensure the identifiability of the parameterizations. A consequence of using identifiable parameterizations was the fact that exponential convergence could be guaranteed under persistency of excitation conditions. It was also shown through an example that exponential convergence could improve robustness to noise and unmodeled dynamics. A MRAC scheme for a MIMO system with some small output noise and some unmodeled dynamics was stable with an identifiable controller, but became unstable with an overparameterized controller.

A second contribution was the derivation of frequency domain conditions on the reference inputs to guarantee parameter convergence in direct adaptive control. In the SISO MRAC case, we recovered the requirement of $2n$ frequencies for $2n$ parameters and, in the SISO PPAC case, the requirement of $4n-2$ frequencies for $4n-2$ parameters. However, in the MIMO MRAC case, the necessary and sufficient conditions were different. When the sufficient condition was violated, but the necessary condition was not, parameter convergence occurred in some cases and not in others. Persistent excitation was not only dependent on the number of frequency components, but also on their location and on the specific system. Furthermore, for the same system, the locations of the frequency components that guaranteed parameter convergence were usually different between MRAC and identification. For the same input frequencies, parameter convergence might occur in MRAC and not in identification, and *vice versa*. One could be tempted to speculate that parameter convergence in the MRAC case did not occur when some frequencies were eliminated (filtered out) by the controller. However, convergence sometimes

occurred in the MRAC and not in the identifier and, clearly, frequencies were not eliminated in that case.

Frequency domain conditions for the pole placement algorithm were also obtained. Again, it was found that the convergence properties were different from the MRAC scheme, or the recursive identifier, and generally more frequency components more required. This can be related to the fact that the pole placement algorithm required as much as twice the number of parameters necessary to estimate in the MRAC algorithm. Since no practical global stability result currently exists without persistency of excitation conditions on the input, the only advantage of PPAC resides in the fact that it can be used with nonminimum phase systems when MRAC cannot.

Another important contribution of this thesis is the modification of MIMO MRAC algorithms in order to guarantee stability when the high-frequency gain matrix is unknown (only an upper bound or a lower bound on the norm of the high-frequency gain matrix and an upper bound on the parameter vector are required). The control parameters are obtained through a particular transformation (a sort of hysteresis) applied to the estimated parameters. The transformation ensures that the matrix of control parameters related to the high-frequency gain matrix remains nonsingular. We showed that all signals in the system remain bounded, that the output error and the regressor error converge to zero. Furthermore, the hysteresis transformation might prove useful in other problems where singularity regions must be avoided such as in pole placement and nonlinear control. Another contribution is the fact that exponential parameter convergence is achieved under persistency of excitation conditions when an identifiable parameterization is used. An example proved that the problem of the singularity of the estimate of the high-frequency gain matrix was more than a mere technicality in a proof of stability (as it has often been considered in the literature). Indeed, a MRAC system was shown to converge with the hysteresis transformation implemented but failed to do so when the transformation was not present.

Using the hysteresis transformation we were able to prove global stability and output error convergence for a MIMO MRAC scheme requiring less *a priori* information (only the relative degrees of the diagonal elements of the Hermite normal form and an upper bound on the maximum of the observability indices).

Finally, we compared pseudo-canonical and canonical input-output models for the identification of multivariable systems. Frequency domain conditions for parameter convergence were discussed, assuming *a priori* knowledge of the structure of the system (the observability indices in the canonical case and pseudo-observability indices in the pseudo-canonical case). It was shown that these conditions were almost identical between pseudo-canonical and canonical models. It was also shown that the *a priori* knowledge of the canonical or pseudo-canonical structure of the system is necessary to define an identifiable controller parameterization in MIMO direct adaptive control. Finally, a contribution was the derivation of criteria for canonical of pseudo-canonical structure identification. If the order of the system is known, identification of pseudo-canonical forms can be more efficient since no information is lost when a

new structure is selected. However, it was shown that if the order is unknown this advantage becomes less important given the larger number of parameters to estimate and given the fact that constraints must be added to the estimation algorithm to guarantee strict properness of the estimate.

6.2 Directions for future work

There exist several research directions. A first research direction consists in reducing further the required *a priori* knowledge about the Hermite normal form. Another research direction is the application of adaptive control techniques to reconfigurable control systems. The study of the robustness properties of multivariable adaptive control schemes and the improvement of these properties are also interesting research topics. Current stability results for MIMO adaptive pole placement schemes are very limited. Therefore, looking for stronger results is a possible research direction. Finally, to apply the singularity avoidance algorithm to other problems requiring a similar mechanism is a promising research topic.

Appendix

Proof of Lemma 2.3

1. We note that $\forall\, t \notin \{\tau_k\}$

$$\frac{dP^{-1}}{dt} = g\frac{\psi\psi^T}{1+\gamma\psi^T\psi} \geq 0$$

Since the right-hand side is bounded, $P^{-1} \in L_{\infty e}$. Further, $P^{-1}(t_1) \geq P^{-1}(t_2)$, $\forall t_1 \geq t_2 \geq 0$ between resettings. At the resettings $P^{-1}(\tau_k) = P^{-1}(0)$. Therefore, $P^{-1}(t) \geq P^{-1}(0) = k_0^{-1}I$ $\forall t \geq 0$, and $0 \leq P(t) \leq k_0 I$. Then, it can also be deduced that $\frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}}$, $\dot{P}$, $\pi = \frac{P\psi}{1+\|\psi_t\|_\infty}$ $\in L_\infty$.

2. The smallest possible difference between τ_{k+1} and τ_k corresponds to $\delta_{\tau_k} = 0$ and a variation of $\lambda_{\max}(P) = \sigma_{\max}(P)$ from k_0 to k_1. Given the continuity of $\sigma_{\max}(P)$ in the elements of P (*cf.* Definition 4.3)

$$(\tau_{k+1} - \tau_k) \geq (k_0 - k_1)\|\dot{P}\|_\infty^{-1}$$

3. If we define the Lyapunov function V as

$$V = \mathrm{tr}(\phi^T P^{-1}\phi) \geq 0$$

then $\forall\, t \notin \{\tau_k\}$

$$\dot{V} = -g\frac{e_2^T e_2}{1+\gamma\psi^T\psi} \leq 0$$

Furthermore, since $P^{-1}(\tau_k) = k_0^{-1}I < k_1^{-1}I \leq P^{-1}(\tau_k^-)$

$$V(\tau_k) < V(\tau_k^-)$$

It follows that

$$0 \leq V(t) \leq V(0) = k_0^{-1}|\phi(0)|^2 \quad \forall t \geq 0$$

and since $P \in L_\infty$, it can be deduced that ϕ, $\dot{\phi}$, β, $\frac{\phi^T\psi}{(1+\gamma\psi\psi^T)^{1/2}} \in L_\infty$.

4.

$$\int_0^\infty \frac{g e_2^T e_2}{1+\gamma\psi^T\psi}d\tau \;=\; V(0) - V(\infty) + \sum_k (V(\tau_k) - V(\tau_k^-)) \;\leq\; V(0)$$

implies that $\frac{e_2}{(1+\gamma\psi^T\psi)^{1/2}} \in L_2$. Furthermore,

$$\dot{\phi} = -g\frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}} \quad \frac{e_2^T}{(1+\gamma\psi^T\psi)^{1/2}}$$

is the product of a term $\in L_\infty$ by a term $\in L_2$. Consequently, $\dot{\phi} \in L_2$. Similarly,

$$\beta = \frac{(1+\gamma\psi^T\psi)^{1/2}}{1+\|\psi_t\|_\infty} \quad \frac{e_2^T}{(1+\gamma\psi^T\psi)^{1/2}}$$

5. Given Definition 2.2, if $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE

$$\forall\, \alpha > 0 \;\; \exists\, t_0 \geq 0 \quad \text{such that} \quad \lambda_{\min}\left(\int_{t_0}^{\infty} \frac{\psi\psi^T}{1+\gamma\psi^T\psi}\, d\tau\right) \;\leq\; \alpha$$

Further,

$$P^{-1}(t) = k_0^{-1} I + \int_{\tau_k}^{t} g\frac{\psi\psi^T}{1+\gamma\psi^T\psi} d\tau \qquad \forall\, \tau_k \leq t < \tau_{k+1}$$

Therefore, if $\alpha = g^{-1}(k_1^{-1} - k_0^{-1})$, then $\exists\, t_0 \geq 0$ such that no more than one covariance resetting may occur after t_0. Consequently, $\{\tau_k\}$ is a finite set.

Since there is a finite number, say N_r, of covariance resettings

$$\int_0^\infty \frac{gP\psi\psi^T P}{1+\gamma\psi^T\psi}\, d\tau \;=\; P(0) - P(\infty) + \sum_{k=1}^{N_r}(P(\tau_k) - P(\tau_k^-)) \;\leq\; (N_r+1)k_0 I$$

Consequently, $\frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}} \in L_2$. Similarly,

$$\dot{P} = -g\frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}} \quad \frac{\psi^T P}{(1+\gamma\psi^T\psi)^{1/2}} \quad \in\; L_2$$

and

$$\pi = \frac{(1+\gamma\psi^T\psi)^{1/2}}{1+\|\psi_t\|_\infty} \quad \frac{P\psi}{(1+\gamma\psi^T\psi)^{1/2}} \quad \in\; L_2$$

Furthermore, since $\dot{P}$ and $\dot{\phi}$ are the products of two terms $\in L_2$, using Schwartz inequality (*cf.* Desoer & Vidyasagar [81], pp. 232), $\dot{P}$ and $\dot{\phi} \in L_1$.

Finally, $\forall$ vector x, define the function $w_x(t) = x^T P(t) x$. Then $\forall\, t \geq \tau_{N_r}$ $(\tau_0 = 0)$

$$\begin{aligned} 0 \;&\leq\; w_x(t) \leq k_0|x|^2 \\ \dot{w}_x(t) \;&=\; -g\frac{x^T P\psi\psi^T P x}{1+\gamma\psi^T\psi} \leq 0 \end{aligned}$$

so that $w_x(t)$ is monotonically decreasing and bounded below $\forall x$. Therefore, $w_x(t)$ converges $\forall x$ and thus $\lim_{t\to\infty} P(t) = P_\infty$. Since $\frac{d}{dt}(P^{-1}\phi) = 0 \quad \forall\, t \notin \{\tau_k\}$, it follows that $\phi(t) = k_0^{-1}P(t)\phi(\tau_{N_r})\ \forall\, t \geq \tau_{N_r}$, and

$$\lim_{t\to\infty} \phi(t) = k_0^{-1}P_\infty\phi(\tau_{N_r}) = \phi_\infty$$

6.a <u>Case $k_1 > 0$:</u>

Since $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE,

$$\forall\, t_0 \geq 0 \;\; \exists\, \delta(t_0) > 0 \quad \text{such that} \quad \lambda_{\min}\Big(\int_{t_0}^{t_0+\delta} \frac{g\psi\psi^T}{1+\gamma\psi^T\psi}\, d\tau\Big) \;\; \geq \;\; k_1^{-1} - k_0^{-1}$$

Therefore, $\forall\, \tau_k \;\exists\, \delta_k > 0$ such that $\lambda_{\max}(P(\tau_k + \delta_k)) = k_1$ and, thus, $\{\tau_k\}$ is an infinite set. Since $\frac{d}{dt}(P^{-1}\phi) = 0 \quad \forall\, t \notin \{\tau_k\}$, it follows that $\phi(t) = k_0^{-1}P(t)\phi(\tau_k)\ \forall\, t$ such that $\tau_k \leq t < \tau_{k+1}$. It follows that

$$|\phi_i(\tau_{k+1})| \;\leq\; \frac{k_1}{k_0}|\phi_i(\tau_k)| \;\leq\; (\frac{k_1}{k_0})^{(k+1)}|\phi_i(0)| \qquad \forall\, i = 1,\ldots,p$$

where ϕ_i is the i-th column of ϕ. Further,

$$|\phi_i(t)| \;\;\leq\;\; |\phi_i(\tau_k)| \;\;\leq\;\; (\frac{k_1}{k_0})^k|\phi_i(0)| \qquad \forall\, t \geq \tau_k \;\; \forall\, i = 1,\ldots,p$$

It follows that $\lim_{t\to\infty} \phi(t) = 0$.

6.b <u>Case $k_1 = 0$:</u>

Since $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE,

$$\lim_{t\to\infty}\; \lambda_{\min}\Big(\int_0^t \frac{g\psi\psi^T}{1+\gamma\psi^T\psi}\, d\tau\Big) \;\; = \;\; \infty$$

Therefore,

$$\lim_{t\to\infty}\; \lambda_{\max}(P(t)) \;\; = \;\; 0$$

Finally, since $\phi(t) = k_0^{-1}P(t)\phi(0)\ \forall\, t \geq 0$, it follows that $\lim_{t\to\infty}\phi(t) = 0$.

7. Since $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is PE,

$$\exists\, \delta > 0 \quad \text{such that} \quad \lambda_{\min}\Big(\int_{t_0}^{t_0+\delta} \frac{g\psi\psi^T}{1+\gamma\psi^T\psi}\, d\tau\Big) \;\; \geq \;\; k_1^{-1} - k_0^{-1} \qquad \forall\, t_0 \geq 0$$

Therefore, $\{\tau_k\}$ is an infinite set. Further,

$$|\phi_i(t)| \;\;\leq\;\; \frac{k_0}{k_1}(\frac{k_1}{k_0})^{t/(\delta+k_2)}|\phi_i(0)| \qquad \forall\, i = 1,\ldots,p$$

Consequently, $\lim_{t\to\infty} \phi(t) = 0$ exponentially.

8. Since $\frac{d}{dt}(P^{-1}\phi) = 0 \ \ \forall\, t \notin \{\tau_k\}$, it follows that $\phi(t) = k_0^{-1}P(t)\phi(\tau_k)\ \forall\, t$ such that $\tau_k \le t < \tau_{k+1}$. It follows that

$$|\phi_i(t)| \le |\phi_i(\tau_k)| \le (\frac{k_1}{k_0})^{(k+1)}|\phi_i(0)| \le |\phi_i(0)| \qquad \forall\, i = 1, \ldots, p$$

where ϕ_i is the i-th column of ϕ. □

Proof of Theorem 2.3

We know from Lemma 2.3 that the columns of the parameter error converge and remain bounded in norm by their initial values. Now, the control law (2.2) can be rewritten

$$u_i = r_i + w_i^T\theta_i \qquad i = 1, \ldots, m$$

where θ_i is the i-th column of θ, assuming the signal vectors w_i are properly defined. If the vector w is defined as

$$w = \left[\ w_1^T \ \ \ldots \ \ w_2^T\ \right]^T$$

then

$$u = r + \tilde{\theta}^T w$$

where $\tilde{\theta}$ = block diag$\{\theta_i\}$. The vector w is made of filtered inputs and outputs. Since $\tilde{\theta}$ is bounded, the signals are all bounded over finite time. Define H_{wu} as the transfer function between w and u, and $w_\infty(t)$, $y_\infty(t)$, and $\psi_\infty(t)$ as the signal vectors w, y_p, and ψ when $\theta = \theta_\infty$, then

$$\begin{aligned} w_\infty &= H_{wu}D_R[(\Lambda - \bar{C}_\infty)D_R - \bar{D}_\infty N_R]^{-1}\Lambda[r] = H_{w_\infty r}[r] \\ y_\infty &= N_R[(\Lambda - \bar{C}_\infty)D_R - \bar{D}_\infty N_R]^{-1}\Lambda[r] = H_{y_\infty r}[r] \\ \psi_\infty &= H_{\psi u}D_R[(\Lambda - \bar{C}_\infty)D_R - \bar{D}_\infty N_R]^{-1}\Lambda[r] = H_{\psi_\infty r}[r] \end{aligned} \tag{7.1}$$

It can be verified that $H_{w_\infty r}$ is stable and proper, and that $H_{y_\infty r}$ and $H_{\psi_\infty r}$ are stable and strictly proper. Furthermore,

$$w = H_{w_\infty r}[r + (\tilde{\theta} - \tilde{\theta}_\infty)^T w]$$

Since θ converges to θ_∞, by the small gain theorem (*cf.* Desoer & Vidyasagar [81]), $w \in L_\infty$. Therefore, $u \in L_\infty$. Finally, since

$$\begin{aligned} y_p &= H_{y_\infty r}[r + (\tilde{\theta} - \tilde{\theta}_\infty)^T w] \\ \psi &= H_{\psi_\infty r}[r + (\tilde{\theta} - \tilde{\theta}_\infty)^T w] \end{aligned}$$

then y_p and $\psi \in L_\infty$. □

Proof of Lemma 3.1

We will prove only the first part of the lemma, since the second part is the dual of the first. Let define $D_i(s)$ and $N_i(s)$ as the i-th row of $D(s)$ and $N(s)$, then by hypothesis (since, from Definition 1.6, the set of observability indices is equal to one of the sets of pseudo-observability indices)

$$D_i(s)N_R(s) = N_i(s)D_R(s) \quad \text{and} \quad \partial c_j D_i \leq \rho_j - 1 \quad \forall j \quad \forall i$$

Suppose that $\{N_L, D_L\}$ is the pseudo-canonical left MFD with indices $\{\rho_i\}$, then (*cf.* Definition 1.12)

$$\partial c_j D_L = \rho_j \quad \partial r_j D_L \leq \max(\rho_j, \rho_{\max} - 1) \qquad \Gamma_c[D_L] = I \qquad D_L(s)N_R(s) = N_L(s)D_R(s)$$

Now, if $\rho_k = \rho_{\max}$, then the matrices $\{N_L^*, D_L^*\}$ defined by

$$\begin{aligned} j\text{th row of } D_L^* &= j\text{th row of } D_L \quad \forall j \neq k \\ k\text{th row of } D_L^* &= k\text{th row of } D_L + D_i \end{aligned}$$

$$\begin{aligned} j\text{th row of } N_L^* &= j\text{th row of } N_L \quad \forall j \neq k \\ k\text{th row of } N_L^* &= k\text{th row of } N_L + N_i \end{aligned}$$

forms another pseudo-canonical left MFD ($\forall i$) with

$$\partial c_j D_L^* = \rho_j \quad \partial r_j D_L^* \leq \max(\rho_j, \rho_{\max} - 1) \qquad \Gamma_c[D_L^*] = I \qquad D_L^*(s)N_R(s) = N_L^*(s)D_R(s)$$

Given the uniqueness of D_L (*cf.* Definition 1.12),

$$D_L^*(s) = D_L(s) \quad \Rightarrow \quad D_i(s) = 0 \quad N_i(s) = 0 \ \forall i \quad \Rightarrow \quad D(s) = 0 \quad N(s) = 0$$

The second part of the lemma can be proved in a similar manner by using a pseudo-canonical right MFD instead of a left one. □

Gram-Schmidt orthogonalization with memory

Let $W(t) \in \Re^{p \times p}$ be a matrix of rank r function of time and let $h > 0$ be a constant to be fixed later.

1. At time $t = 0$, apply the following procedure:

$$Y_1 = W_{k^{(1)}} \quad \text{where} \quad \begin{array}{l} W_{k^{(1)}} = k^{(1)}\text{th column of } W \text{ s.t.} \\ |W_{k^{(1)}}| = \max_k |W_k| \text{ and } k^{(1)} \text{ minimum} \end{array}$$

$$
\begin{aligned}
W^{(1)} &= (I - \frac{Y_1 Y_1^T}{Y_1^T Y_1})W \qquad (W^{(1)}_{k^{(1)}} = 0) \\
Y_2 &= W^{(1)}_{k^{(2)}} \quad \text{where} \quad \begin{array}{l} W^{(1)}_{k^{(2)}} = k^{(2)}\text{th column of } W^{(1)} \text{ s.t.} \\ |W^{(1)}_{k^{(2)}}| = \max_k |W^{(1)}_k| \text{ and } k^{(2)} \text{ minimum} \end{array} \\
W^{(2)} &= (I - \frac{Y_2 Y_2^T}{Y_2^T Y_2})W^{(1)} \qquad (W^{(2)}_{k^{(1)}} = W^{(2)}_{k^{(2)}} = 0) \\
&\vdots \\
Y_r &= W^{(r-1)}_{k^{(r)}} \quad \text{where} \quad \begin{array}{l} W^{(r-1)}_{k^{(r)}} = k^{(r)}\text{th column of } W^{(r-1)} \text{ s.t.} \\ |W^{(r-1)}_{k^{(r)}}| = \max_k |W^{(r-1)}_k| \text{ and } k^{(r)} \text{ minimum} \end{array} \\
W^{(r)} &= (I - \frac{Y_r Y_r^T}{Y_r^T Y_r})W^{(r-1)} = 0
\end{aligned}
$$

Then, the matrix

$$X(0) = \left[\frac{Y_1}{|Y_1|} \quad \cdots \quad \frac{Y_r}{|Y_r|} \right]$$

is an orthogonal basis of $\mathcal{R}(W(0))$, the space of dimension r spanned by the columns of $W(0)$. The order of selection of the columns of $W(0)$, $\{k^{(i)}\}_{|t=0}$, $i = 1, \ldots, r$ is uniquely defined by this procedure. Therefore, the matrix $X(0)$ is also uniquely defined.

2. Keep the initial order of selection of the columns, $\{k^{(i)}\}_{|t=0}$, for the orthogonalization of $W(t)$, $t > 0$, until

$$\exists\, t = t_1(W) > 0 \quad \exists\, 1 \le j \le r \quad \exists\, 1 \le l \le p \quad s.t. \quad |W^{(j-1)}_{k(j)}(t)|^2 + h < |W^{(j-1)}_l(t)|^2$$

where $W^{(0)} = W$. Then, at $t = t_1(W)$ the procedure defined for $t = 0$ is applied for the computation of $X(t_1)$, and a new order of selection of the columns, $\{k^{(i)}\}_{|t=t_1(W)}$, is found.

3. Finally, continue this procedure for the orthogonalization of $W(t)$, $t > t_1(W)$. The set $\{t_k(W)\}$ are the time instants when the order of selection of the columns of W is changed. The matrix $X(t)$ will be uniquely defined for all $t \ge 0$.

The procedure is applied with $W = P_X$, leading to $X = V_f$, and similarly to P_Y, leading to U_0. The advantage of using this Gram-Schmidt orthogonalization with memory is that the matrices V_f and U_0 are uniquely defined for given matrices P_X and P_Y. The constant h is chosen sufficiently small that no vector Y_i may become equal to zero, while preventing the order of selection of the columns to change an infinite number of time during a finite time interval. Since the order of selection of the columns is kept constant for some interval of time, V_f and U_0 are piecewise continuous (the continuity property is discussed in greater detail in the proof

of Lemma 4.2 in the appendix). Finally, note that using the Gram-Schmidt orthogonalization procedure with memory for $P_{\mathcal{X}} = I$ or $P_{\mathcal{Y}} = I$ leads to $V_f = I$ and $U_0 = I$.

We will now show that if

$$h \leq \frac{1}{2p}$$

the Gram-Schmidt orthogonalization with memory is always well-defined for V_f and U_0. In other words, we show that $|Y_i(t)| > 0 \; \forall \, t \geq 0$.

Indeed, let $W = P_{\mathcal{X}}$. The Gram-Schmidt orthogonalization procedure with memory is applied to W at time $t = 0$ and we let $\{k^{(i)}\}$, $i = 1, \ldots, N_P$ be the order of selection of the columns of $W(0)$. Given the properties of projections (*cf.* Definition 4.2), as long as $|Y_i| = |W_{k(i)}^{(i-1)}| > 0$

$$\begin{aligned}
W^{(0)} &= W = V_f V_f^T = \left[\frac{Y_1}{|Y_1|} \; \cdots \; \frac{Y_{N_P}}{|Y_{N_P}|} \right] \left[\frac{Y_1}{|Y_1|} \; \cdots \; \frac{Y_{N_P}}{|Y_{N_P}|} \right]^T \\
W^{(i)} &= \left[0 \; \ldots \; 0 \;\; \frac{Y_{i+1}}{|Y_{i+1}|} \; \cdots \; \frac{Y_{N_P}}{|Y_{N_P}|} \right] \left[\frac{Y_1}{|Y_1|} \; \cdots \; \frac{Y_{N_P}}{|Y_{N_P}|} \right]^T \qquad i = 1, \ldots, N_P - 1 \\
&= \left[\frac{Y_{i+1}}{|Y_{i+1}|} \; \cdots \; \frac{Y_{N_P}}{|Y_{N_P}|} \right] \left[\frac{Y_{i+1}}{|Y_{i+1}|} \; \cdots \; \frac{Y_{N_P}}{|Y_{N_P}|} \right]^T \qquad i = 1, \ldots, N_P - 1
\end{aligned}$$

so that $\|W^{(i)}\|_F^2 = N_P - i$. Note that $W^{(i)}$ has i columns equal to zero, when $i = 0, \ldots, N_P - 1$, and $W^{(N_P)} = 0$. Therefore,

$$\begin{aligned}
\max_k |W_k^{(i)}|^2 &\geq \frac{N_P - i}{p - i} \qquad i = 0, \ldots, N_P - 1 \\
&\geq \frac{1}{p - N_P + 1}
\end{aligned}$$

So, if $h \leq \frac{1}{2(p - N_P + 1)}$ then, whenever the order is preserved

$$|W_{k(i+1)}^{(i)}|^2 \geq \max_k |W_k^{(i)}|^2 - h \geq \frac{1}{2(p - N_P + 1} \qquad i = 0, \ldots, N_P - 1$$

Consequently, $|Y_i(t)|^2 \geq \frac{1}{2(p - N_P + 1)} \; \forall \, i = 1, \ldots, N_P, \; \forall \, t \geq 0$. Similarly, if $h \leq \frac{1}{2(p - N_P + 1)}$ then the orthogonalization of $P_{\mathcal{Y}}$ is well-defined $\forall \, t \geq 0$. Finally, since $N_P \geq 1$, the result follows. □

Proof of Lemma 4.2

1. Given the properties of the singular value decomposition (*cf.* Definition 4.3), $V_P \Sigma_P^{-1} U_P^T = \sum_{i=1}^{N_P} \frac{1}{\sigma_i} V_{P_i} U_{P_i}^T$ is uniquely defined. Similarly, $U_P U_P^T$ and $U_\epsilon U_\epsilon^T = I - U_P U_P^T$ are uniquely defined. Given the definition of the Gram-Schmidt orthogonalization with memory, V_f and U_0 are uniquely defined. Therefore, Q is uniquely defined.

2. If $T_{k-1} \leq t < t_k$ or if $t_k \leq t < T_k$ and $\sigma_{\min}(P_p^T(t)) > \epsilon_P$, then $C_{0_c}^T(t) = C_{0_f}^T(t)$ and $\sigma_{\min}(C_{0_c}(t)) \geq \sigma/b$. If $t_k \leq t < T_k$ and $\sigma_{\min}(P_p^T) \leq \epsilon_P$, then

$$C_{0_c}(t) = C_{0_c}(t_{ij}^k) + (C_0(t) - C_0(t_{ij}^k)) + (P_p^T(t) - P_p^T(t_{ij}^k))Q(t_{ij}^k) \tag{7.2}$$

where, from (4.18) and (4.19) with (4.16) and (4.17), $\sigma_{\min}(C_{0_c}(t_{ij}^k)) \geq \sigma/b$. Using (2.29) and condition (4.20) with (7.2)

$$\sigma_{\min}(C_{0_c}(t)) \geq \sigma/b - \delta_t(\|\dot{C}_0\|_\infty + \|\dot{P}_p\|_\infty \frac{2\|C_0\|_\infty}{\epsilon_P + \delta_P}) > (\sigma/b)(1 - 1/c)$$

Consequently, $\sigma_{\min}(C_{0_c}(t)) > (\sigma/b)(1 - 1/c) > 0 \;\; \forall\, t > 0$.

3. If $\sigma_{\min}(P_p^T) > \epsilon_P$, then $(P_p^T P_p)^{-1} < 1/\epsilon_P^2 I$ and $Q \in L_\infty$ since P, C_0, and $C_{0_f} \in L_\infty$.
If $\sigma_{\min}(P_p^T) \leq \epsilon_P$ and $\sigma_{\max}(P_p^T) > \epsilon_P + (p-1)\delta_P$, then $Q \in L_\infty$ since $\Sigma_P^{-1} < 1/(\epsilon_P + \delta_P)$.
If $\sigma_{\min}(P_p^T) \leq \epsilon_P$ and $\sigma_{\max}(P_p^T) \leq \epsilon_P + (p-1)\delta_P$, since

$$C_0^T(t) = C_0^{*^T} + P_p^T(t)\frac{\phi(\tau_k)}{k_0} \qquad \forall\, \tau_k \leq t < \tau_{k+1}$$

then

$$\sigma_{\min}(C_0) \geq \sigma_{\min}(C_0^*) - (\epsilon_P + (p-1)\delta_P)\frac{|\phi(0)|}{k_0} > \sigma/a$$

and therefore, $Q = 0$. It follows that $Q \in L_\infty$, $\theta_c \in L_\infty$, and $\phi_c \in L_\infty$.

4.

$$\beta_c = \frac{\phi_c^T \psi}{1 + \|\psi_t\|_\infty} = \frac{\phi^T \psi}{1 + \|\psi_t\|_\infty} + \frac{Q^T P \psi}{1 + \|\psi_t\|_\infty} = \beta + Q^T \pi$$

From Lemma 2.3, $\beta \in L_2 \cap L_\infty$. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, then $\pi \in L_2 \cap L_\infty$, and $\beta_c \in L_2 \cap L_\infty$. If $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, then $\lim_{t\to\infty} C_0(t) = C_0^*$. Therefore, $\exists\, T > 0$ such that $|\sigma_{\min}(C_0) - \sigma_{\min}(C_0^*)| < \sigma(1 - 1/a)$, $\forall\, t > T$ and $\beta_c(t) = \beta(t)\ \forall\, t > T$. Consequently, $\beta_c \in L_2 \cap L_\infty$.

5. If $\lim_{t\to\infty} \phi(t) = 0$ then $\lim_{t\to\infty} C_0(t) = C_0^*$ and $\exists\, T > 0$ such that $|\sigma_{\min}(C_0) - \sigma_{\min}(C_0^*)| < \sigma(1 - 1/a)$, $\forall\, t > T$ and $\phi_c(t) = \phi(t)\ \forall\, t > T$.

6. The smallest difference between t_k and T_k or T_k and t_{k+1} corresponds to a variation of $\sigma_{\min}(C_0)$ from σ/b to σ/a or σ/a to σ/b. Given the continuity of the singular values of C_0 in the elements of C_0 (*cf.* Definition 4.3)

$$\begin{aligned} (T_k - t_k) &\geq \sigma(1/a - 1/b)\|\dot{C}_0\|_\infty^{-1} \\ (t_k - T_{k-1}) &\geq \sigma(1/a - 1/b)\|\dot{C}_0\|_\infty^{-1} \end{aligned}$$

7. From Lemma 2.3, C_0 is a matrix of continuous functions of time and $\exists$ a matrix C_∞ such that $\lim_{t\to\infty} C_0 = C_\infty$. Consequently, given the continuity of the singular values, $\forall \epsilon > 0$, $\exists T > 0$

such that $\forall t > T$, $|\sigma_{\min}(C_0) - \sigma_{\min}(C_\infty)| < \epsilon$. Now, suppose that $\sigma_{\min}(C_\infty) \neq \sigma/a$, then $\exists \epsilon > 0$ such that $|\sigma_{\min}(C_\infty) - \sigma/a| > \epsilon$ and $\exists T > 0$ such that $\forall t > T$, $|\sigma_{\min}(C_0) - \sigma_{\min}(C_\infty)| < \epsilon$. Therefore, no more jumps can occur after T sufficiently large. If $\sigma_{\min}(C_\infty) = \sigma/a$, then $\exists T > 0$ such that $\forall t > T$, $|\sigma_{\min}(C_0) - \sigma_{\min}(C_\infty)| < \sigma(1/a - 1/b)$ and, therefore, no more than one jump may occur $\forall t > T$. Consequently, $\{T_k\}$ and $\{t_k\}$ are finite sets.

8. $P(t)$ is symmetric, positive definite, and monotonically decreasing between the covariance resetting time instants, $\{\tau_k\}$. In other words,

$$P(t_1) \geq P(t_2) \qquad \forall\, \tau_k < t_1 < t_2 < \tau_{k+1}$$

Therefore,

$$P(t_1)P(t_1) \geq P(t_2)P(t_2) \qquad \forall\, \tau_k < t_1 < t_2 < \tau_{k+1}$$

and

$$P_p^T(t_1)P_p(t_1) \geq P_p^T(t_2)P_p(t_2) \qquad \forall\, \tau_k < t_1 < t_2 < \tau_{k+1}$$

Consequently, from Theorem 5.10, pp. 315 in Stewart (1973)

$$\sigma_i(P_p^T(t_1)) \geq \sigma_i(P_p^T(t_2)) \qquad \forall\, i \quad \forall\, \tau_k < t_1 < t_2 < \tau_{k+1} \tag{7.3}$$

and there cannot exist more than one time instant when $\sigma_{\min}(P_p^T)$ becomes smaller than ϵ_P in between two resetting time instants. Therefore, the set $\{t_{0j}^k\}$ is a finite set, since, either $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE and $\{\tau_k\}$ is a finite set (*cf.* Lemma 2.3), or $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is SE, $\lim_{t\to\infty} C_0(t) = C_0^*$, and $\exists\, T > 0$ such that $Q(t) = 0\ \forall\, t > T$. However, $\{t_{ij}^k\}$ could be an infinite set, when $\frac{\psi}{(1+\gamma\psi^T\psi)^{1/2}}$ is not SE, since i could be infinite.

9. If we can show that $U_P\Sigma_P^{-1}V_P^T$ and R_P converge when P and C_0 converge, then we will have proved that $\phi_c(t)$ converges to some $\phi_{c\infty}$ when $P(t)$ and $\phi(t)$ converge to some P_∞ and ϕ_∞.

Part 1: $U_P\Sigma_P^{-1}V_P^T$ converges

Let $\{\sigma_i\}$ be the singular values of P_p^T, then

$$U_P\Sigma_P^{-1}V_P^T = \sum_{i=1}^{N_P} \frac{1}{\sigma_i^2} U_{P_i}U_{P_i}^T P_p^T$$

where U_{P_i} is the i-th column of U_P. Therefore, assuming that N_P stays constant

$$\begin{aligned}
\delta(U_P\Sigma_P^{-1}V_P^T) &= \sum_{i=1}^{N_P} \frac{1}{(\sigma_i + \delta\sigma_i)^2}(U_{P_i}U_{P_i}^T + \delta(U_{P_i}U_{P_i}^T))(P_p^T + \delta P_p^T) - \frac{1}{\sigma_i^2}U_{P_i}U_{P_i}^T P_p^T \\
&= \sum_{i=1}^{N_P} \frac{-(2\sigma_i + \delta\sigma_i)\delta\sigma_i}{(\sigma_i + \delta\sigma_i)^2\sigma_i^2} U_{P_i}U_{P_i}^T P_p^T + \frac{1}{(\sigma_i + \delta\sigma_i)^2} U_{P_i}U_{P_i}^T \delta P_p^T \\
&\quad + \frac{1}{(\sigma_i + \delta\sigma_i)^2}\delta(U_{P_i}U_{P_i}^T)(P_p^T + \delta P_p^T)
\end{aligned} \tag{7.4}$$

Now, $\forall\ \Delta \leq \delta_P/2$, we can define a unique subset of distinct singular values of Σ_P, say $\{\sigma_{s_i}\}$, such that

$$\sigma_{s_1} > \sigma_{s_2} + 2\Delta > \sigma_{s_3} + 4\Delta > \ldots > \sigma_{s_q} + 2(q-1)\Delta \quad \text{with} \quad s_q = N_P$$

$$\begin{array}{l}\sigma_{s_{(i-1)}} - 2\Delta \ \geq \ \sigma_{s_{(i-1)}+1} \leq \sigma_{s_{(i-1)}+2} + 2\Delta \leq \ldots \leq \sigma_{s_i} + 2(s_i - s_{(i-1)} - 1)\Delta \\ \quad i = 1, \ldots, q \quad \text{with} \quad s_0 = 0\end{array} \tag{7.5}$$

where q is the size of the set $\{\sigma_{s_i}\}$, *i.e.*, $1 \leq q \leq N_P$.

Let $m_i = s_i - s_{i-1}$ for $i = 1, \ldots, q$ with $\sum_{i=1}^{q} m_i = N_P$. Basically, we have divided the N_P singular values of Σ_P in q clusters of m_i singular values, $i = 1, \ldots, q$, such that the maximum distance between a singular value and its nearest neighbor inside a cluster is smaller than 2Δ but the minimum distance between two clusters is greater than 2Δ. Then, (7.4) can be rewritten in the following manner

$$\begin{aligned}\delta(U_P\Sigma_P^{-1}V_P^T) \ &= \ \sum_{i=1}^{N_P} \frac{-(2\sigma_i + \delta\sigma_i)\delta\sigma_i}{(\sigma_i + \delta\sigma_i)^2\sigma_i^2} U_{P_i} U_{P_i}^T P_p^T + \frac{1}{(\sigma_i + \delta\sigma_i)^2} U_{P_i} U_{P_i}^T \delta P_p^T \\ &+ \sum_{i=1}^{q} \frac{1}{(\sigma_{s_i} + \delta\sigma_{s_i})^2} \sum_{k=1}^{m_i} \delta(U_{P_{s_{(i-1)}+k}} U^T_{P_{s_{(i-1)}+k}})(P_p^T + \delta P_p^T) \\ &+ \sum_{i=1}^{q} \sum_{k=1}^{m_i-1} \frac{(\sigma_{s_i} + \delta\sigma_{s_i})^2 - (\sigma_{s_{(i-1)}+k} + \delta\sigma_{s_{(i-1)}+k})^2}{(\sigma_{s_i} + \delta\sigma_{s_i})^2(\sigma_{s_{(i-1)}+k} + \delta\sigma_{s_{(i-1)}+k})^2} \delta(U_{P_{s_{(i-1)}+k}} U^T_{P_{s_{(i-1)}+k}})(P_p^T + \delta P_p^T)\end{aligned} \tag{7.6}$$

From Definition 4.3, $|\delta\sigma_i| \leq |\delta P_p^T|$. Also, Σ_P and $\Sigma_P + \delta\Sigma_P$ are $\geq (\epsilon_P + \delta_P)I$, and $|P_p^T| \leq k_0$. Furthermore, using Lemma 4.1.1 with (7.5), $\forall\ \Delta \leq \delta_P/2$, if $|\delta P_p^T| < \Delta$ then

$$\|\sum_{k=1}^{m_i} \delta(U_{P_{s_{(i-1)}+k}} U^T_{P_{s_{(i-1)}+k}})\|_F \ \leq \ \frac{2\sqrt{m_i}}{\Delta}\|\delta P_p^T\|_F \qquad i = 1, \ldots, q$$

Therefore, $\forall\ \Delta \leq \delta_P/2$, if $|\delta P_p^T| < \Delta$, then (7.6) yields

$$\begin{aligned}|\delta(U_P\Sigma_P^{-1}V_P^T)| \ &\leq \ \frac{2k_0N_P}{(\epsilon_P + \delta_P)^3}|\delta P_p^T| + \frac{N_P}{(\epsilon_P + \delta_P)^2}|\delta P_p^T| + \frac{2k_0 p \sum_{i=1}^{q}\sqrt{m_i}}{(\epsilon_P + \delta_P)^2}\frac{|\delta P_p^T|}{\Delta} \\ &\quad + \frac{4k_0}{(\epsilon_P + \delta_P)^3}((N_P - q)|\delta P_p^T| + (\sum_{i=1}^{q}\sum_{k=1}^{m_i-1} k)\Delta) \\ &\leq \ \alpha_1|\delta P_p^T| + \alpha_2\frac{|\delta P_p^T|}{\Delta} + \alpha_3\Delta\end{aligned}$$

if α_1, α_2, and α_3 are properly defined. Note also that if $|\delta P_p^T| < \min(1, \delta_P^2/4)$, then

$$|\delta(U_P\Sigma_P^{-1}V_P^T)| \leq \alpha_1|\delta P_p^T| + (\alpha_2 + \alpha_3)|\delta P_p^T|^{1/2} \tag{7.7}$$

Now, $\forall\, \epsilon > 0$, we can choose $\Delta = \min(\frac{\delta_P}{2}, \frac{\epsilon}{2\alpha_3})$. Then,

$$\forall\, \epsilon > 0, \quad \exists\, \delta = \min(\frac{\epsilon^2}{2(\alpha_1\epsilon + 2\alpha_2\alpha_3)}, \frac{\epsilon\delta_P}{2(\alpha_1\delta_P + 2\alpha_2)}, \frac{\epsilon}{2\alpha_3}, \frac{\delta_P}{2}), \quad \text{such that}$$
$$|\delta P_p^T| < \delta \quad \Rightarrow \quad |\delta(U_P \Sigma_P^{-1} V_P^T|) < \epsilon$$

Therefore, $U_P \Sigma_P^{-1} V_P^T$ is continuous in the elements of P_p^T as long as N_P stays constant. From (7.3)

$$N_P(t_1) \geq N_P(t_2) \qquad \forall\, \tau_k < t_1 < t_2 < \tau_{k+1}$$

Further, $N_P \geq 1$. Consequently, if P converges then N_P converges and $U_P \Sigma_P^{-1} V_P^T$ will converge.

Part 2: R_P converges

Given the definition of U_P, $\sigma_{\min}(\Sigma_P) > \sigma_{\max}(\Sigma_\epsilon) + \delta_P$. If $|\delta P_p^T| < \delta_P/2$, using Lemma 4.1.2

$$|\delta(U_P U_P^T)| < \frac{2}{\delta_P} |\delta P_p^T| \tag{7.8}$$

Therefore, $U_P U_P^T$ is continuous in the elements of P_p^T as long as N_P is constant. The same is true for $U_\epsilon U_\epsilon^T$ since $U_\epsilon U_\epsilon^T = I - U_P U_P^T$.

The singular values of $C_{0_f} U_P U_P^T$ have the following properties

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq (\sigma_{N_P} \geq \sigma/b) > \sigma_{N_P+1} = 0$$

Since $\mathcal{X} = \mathcal{R}(C_{0_f} U_P U_P^T)$, using Lemma 4.1.2 and (7.8), if $|\delta P_p^T| < \delta_P/2$ then

$$|\delta P_{\mathcal{X}}| < \frac{2}{\delta_P} \frac{b\|C_0\|_\infty}{\sigma} |\delta P_p^T|$$

The singular values of $C_0 U_\epsilon U_\epsilon^T$ have the following properties

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq (\sigma_{p-N_P} \geq \sigma/a) > \sigma_{p-N_P+1} = 0$$

Since $\mathcal{Y} = \mathcal{R}(C_0 U_\epsilon U_\epsilon^T)^\perp$, using Lemma 4.1.2, (7.8), and the fact that $U_\epsilon U_\epsilon^T = I - U_P U_P^T$, if $|\delta P_p^T| < \delta_P/2$ then

$$|\delta P_{\mathcal{Y}}| < \frac{2}{\delta_P} \frac{a\|C_0\|_\infty}{\sigma} |\delta P_p^T| + \frac{a}{\sigma} |\delta C_0|$$

When the order of selection of the columns stays constant in the Gram-Schmidt orthogonalization procedure with memory, it can be verified that if $|\delta P_p^T| < \delta_P/2$ then

$$\exists\, \eta_1 > 0 \quad \text{s.t.} \quad |\delta V_f| \; < \; \eta_1 |\delta P_{\mathcal{X}}| \; < \; \frac{2\eta_1}{\delta_P} \frac{b\|C_0\|_\infty}{\sigma} |\delta P_p^T| \tag{7.9}$$

$$\exists\, \eta_2 > 0 \quad \text{s.t.} \quad |\delta U_0| \; < \; \eta_2 |\delta P_{\mathcal{Y}}| \; < \; \frac{2\eta_2}{\delta_P} \frac{a\|C_0\|_\infty}{\sigma} |\delta P_p^T| + \eta_2 \frac{a}{\sigma} |\delta C_0| \tag{7.10}$$

Therefore, as long as the order of selection of the columns stays constant V_f is continuous in the elements of $P_{\mathcal{X}}$ and U_0 is continuous in the elements of $P_{\mathcal{Y}}$. Furthermore, since $\dot{P}_p^T$ and $\dot{C}_0$ are bounded, because of the constant $h > 0$, the order of selection of the columns will always remain constant for some minimum amount of time. Therefore, as long as N_P is constant and the order of selection of the columns is constant in the Gram-Schmidt orthogonalization, R_P is continuous in the elements of P_p^T and C_0. Finally, if P_p^T and C_0 converge, because of the continuity property, there will be only a finite number of changes in the order of selection of the columns for the orthogonalization procedure, *i.e.*, $\{t_k(P_{\mathcal{X}})\}$ and $\{t_k(P_{\mathcal{Y}})\}$ will be finite sets and R_P will also converge. □

Proof of existence and uniqueness of the solution

Assume that $\{A_P, B_P, C_P\}$ is a minimal state-space representation of the plant

$$P(s) = C_P(sI - A_P)^{-1}B_P$$

Define $\{A_\lambda, B_\lambda, C_\lambda\}$ as a minimal state-space representation of the controller such that

$$C_\lambda(sI - A_\lambda)^{-1}B_\lambda = \frac{1}{\lambda(s)}\begin{bmatrix} I_p \\ sI_p \\ \vdots \\ s^{\nu-2}I_p \end{bmatrix}$$

Then, the plant with observers $w^{(1)}$, $w^{(2)}$ is described by

$$\begin{aligned} \dot{x}_P &= A_P x_P + B_P u \\ \dot{x}_\lambda^{(1)} &= A_\lambda x_\lambda + B_\lambda u \\ \dot{x}_\lambda^{(2)} &= A_\lambda x_\lambda + B_\lambda y_p \end{aligned}$$

$$\begin{aligned} y_p &= C_P x_P \\ w^{(1)} &= C_\lambda x_\lambda^{(1)} \\ w^{(2)} &= C_\lambda x_\lambda^{(2)} \end{aligned}$$

Define $x_{P\lambda}$ as $\begin{bmatrix} x_P^T & x_\lambda^{(1)^T} & x_\lambda^{(2)^T} \end{bmatrix}^T$, then by using (4.21) and by properly defining the constant matrices $\{A_m, B_m, C_m\}$ (*cf.* Sastry & Bodson [3], section 3.5, pp. 135-136)

$$\begin{aligned} \dot{x}_{P\lambda} &= A_m x_{P\lambda} + B_m \phi_c^T w + B_m C_0^* r \\ y_p &= C_m x_{P\lambda} \end{aligned}$$

Since, model matching must occur when $\phi_c = 0$

$$C_m(sI - A_m)^{-1}B_m = H(s)C_0^{*^{-1}}$$

and the model and its output can be represented by

$$\begin{aligned} \dot{x}_m &= A_m x_m + B_m C_0^* r \\ y_m &= C_m x_m \end{aligned}$$

The representation is non-minimal, but it can be shown (*cf.* Sastry & Bodson [3], section 3.5, pp. 136-137), that the additional modes of A_m are those of N_R and Λ, stable by assumptions.

Now, let $e = x_{P\lambda} - x_m$, then, if the constant matrix R_m is properly defined (*cf.* Sastry & Bodson [3], section 3.5, pp. 137-138)

$$\begin{aligned} \dot{e} &= A_m e + B_m \phi_c^T w_m + B_m \phi_c^T R_m e \\ e_0 &= C_m e \end{aligned}$$

where $w_m = R_m x_m = R_m(sI - A_m)^{-1}B_m C_0^* r$, is a bounded continuous exogenous input. Finally, if $\{A_\psi, B_\psi, C_\psi, D_\psi\}$ are properly defined

$$\begin{aligned} \dot{x}_\psi &= A_\psi x_\psi + B_\psi x_{P\lambda} \\ \psi &= C_\psi x_\psi + D_\psi y_p \end{aligned}$$

where the modes of A_ψ are the poles of $l(s)$. Let $e_r = x_\psi - x_{\psi_m}$, where x_{ψ_m} are the corresponding model signals, then the overall adaptive system has the following form

$$\begin{aligned} \dot{e} &= A_m e + B_m \phi_c^T w_m + B_m \phi_c^T R_m e \\ \dot{e}_r &= A_\psi e_r + B_\psi e \\ \dot{\phi} &= -g \frac{P(e_\psi + \psi_m)(e_\psi + \psi_m)^T \phi}{1 + \gamma (e_\psi + \psi_m)^T (e_\psi + \psi_m)} \\ \dot{P} &= -g \frac{P(e_\psi + \psi_m)(e_\psi + \psi_m)^T P}{1 + \gamma (e_\psi + \psi_m)^T (e_\psi + \psi_m)} \qquad P(0) = P(\tau_k^+) = k_0 I \\ \phi_c &= \phi + PQ \\ e_0 &= C_m e \\ e_\psi &= C_\psi e_r + D_\psi e_0 \end{aligned} \tag{7.11}$$

where $\psi_m = [C_\psi(sI - A_\psi)^{-1}B_\psi + D_\psi C_m](sI - A_m)^{-1}B_m C_0^* r$ is a bounded continuous exogenous input. The matrix Q is a piecewise continuous function of time and is continuous in the elements of ϕ and P in between discontinuities. Therefore, the overall adaptive system can be put under the form of (4.29) with $f(t, x)$ piecewise continuous in t.

Now, we can use Lemma 4.3 to prove existence and uniqueness. First, we will assume for simplicity that $k_1 = 0$ (no covariance resettings). Let $x_0 = x(0)$ and assume that $Q = 0 \; \forall \; t > 0$, then from (7.11), $f(t, x)$ is continuous in t and x, and

$$\begin{aligned} |f(t,x_0)| &\leq (|A_m| + |B_m||R_m||\phi(0)| + |B_\psi|)|e(0)| + |A_\psi||e_r(0)| \\ + & \quad (|B_m|\|w_m\|_\infty + gk_0 \max(1,\gamma^{-1}))|\phi(0)| + gk_0^2 \max(1,\gamma^{-1}) \\ &\leq k_1(|e(0)|^2 + |e_r(0)|^2)^{1/2} + k_2 = h_0 \end{aligned}$$

if k_1 and k_2 are properly defined positive constants. Further, let $r_0 > 0$ finite, then $\forall \; x \in \mathcal{B}_{r_0}(x_0)$ and $\forall \; t \geq 0$ it can be checked that $f(t, x)$ in (7.11) is locally Lipschitz:

$$\begin{aligned} |f(t,x_1) - f(t,x_2)| \leq \; & (|A_m| + |B_m||R_m||\phi_1| + |B_\psi| + g(1+\min(\gamma,1))|D_\psi||C_m|[\min(1,\gamma^{-1}) \\ & +|D_\psi||C_m||e_2| + |C_\psi||e_{r_2}| + \|\psi_m\|_\infty](|\phi_2| + |P_2|)|P_2|)|e_1 - e_2| \\ & +(|A_\psi| + g(1+\min(\gamma,1))|C_\psi|[\min(1,\gamma^{-1}) + |D_\psi||C_m||e_2| + |C_\psi||e_{r_2}| \\ & +\|\psi_m\|_\infty](|\phi_2| + |P_2|)|P_2|)|e_{r_1} - e_{r_2}| \\ & +(|B_m||R_m||e_2| + g\min(1,\gamma^{-1})|P_2|)|\phi_1 - \phi_2| \\ & +g\min(1,\gamma^{-1})(|\phi_1| + |P_1| + |P_2|)|P_1 - P_2| \\ \leq \; & (k_3(|e(0)|^2 + |e_r(0)|^2)^{1/2} + k_4)|x_1 - x_2| = l_0|x_1 - x_2| \end{aligned} \tag{7.12}$$

if k_3 and k_4 are properly defined positive constants. Also,

$$\frac{\max(k_3,k_4)}{k_1+k_2} = k_5 \leq \frac{l_0}{h_0} \leq \frac{k_3+k_4}{\max(k_1,k_2)} = k_6$$

Therefore, from Lemma 4.3, if

$$\delta_0 = \frac{1}{l_0}\min(\rho, \frac{r_0}{k_6^{-1}+r_0}, x) = \frac{k_7}{l_0} \quad \text{with} \quad 0 < \rho < 1 \quad x\exp(x) = k_5 r_0$$

then (7.11) has a unique solution on $[0, \delta_0]$. From Lemma 2.3, we know that as long as a solution exists, then $|\phi(t)| \leq |\phi(0)|$ and $|P(t)| \leq k_0$. Therefore, $|\phi(\delta_0)| \leq |\phi(0)|$, $|P(\delta_0)| \leq k_0$, and $(|e(\delta_0)|^2 + |e_r(\delta_0)|^2)^{1/2} \leq (|e(0)|^2 + |e_r(0)|^2)^{1/2} + r_0$. Then, it can be verified that $\forall \; t \geq \delta_0$,

$$\begin{aligned} |f(t,x_{\delta_0})| &\leq k_1(|e(\delta_0)|^2 + |e_r(\delta_0)|^2)^{1/2} + k_2 = h_1 \\ |f(t,x_1) - f(t,x_2)| &\leq (k_3(|e(\delta_0)|^2 + |e_r(\delta_0)|^2)^{1/2} + k_4)|x_1 - x_2| = l_1|x_1 - x_2| \\ &\leq (l_0 + k_3 r_0)|x_1 - x_2| \quad \forall \; x_1, x_2 \in \mathcal{B}_{r_0}(x(\delta_0)) \end{aligned}$$

Therefore, if $\delta_1 = \frac{k_7}{l_0+k_3r_0}$, then (7.11) has a unique solution on $[\delta_1, \delta_0]$. We can keep extending the solution in this manner indefinitely since $\lim_{t\to\infty}\sum_{i=0}\frac{k_7}{l_0+ik_3r_0}$ is unbounded. However, at $t = t_1$, Q changes, $Q = P_p(P_p^T P_p)^{-1}(C_{0_f}^T - C_0)$ as long as $\sigma_{\min}(P_p^T) > \epsilon_P$. Then, it can be proved,

similarly to the case $Q = 0$, that there exists a unique solution $\forall\, t \in [t_1, t_{01}^1]$. This reasoning can be easily extended for all the subsequent time intervals and, also, when there is some covariance resetting ($k_1 > 0$). Since the time intervals between the time instants $\{\tau_k\}$, $\{t_k\}$, $\{T_k\}$, and $\{t_{ij}^k\}$ are always bounded below by δ_t (if we assume that the covariance resetting are synchronized with the time instants $\{t_{ij}^k\}$ when necessary ($k_2 = \delta_t$)), the interval of existence of the solution can be extended indefinitely and the system has a unique solution $\forall\, t > 0$. □

Proof of Lemma 4.7

The proof is similar in its construction to the proof of Lemma C4 in Ioannou & Tsakalis [90]. From Lemma 4.5, $u \in L_\infty$. Now, assume that $L(s) = \text{diag}\{l(s)\}$, with $l(s)$ polynomial, stable, $\partial l(s) \geq d$, where d is the maximum relative degree, (degree of numerator - degree of denominator), of all elements of $H^{-1}(s)$. Then, using Lemma 4.6, $L^{-1}[u] = (HL)^{-1}[y] \in L_\infty$ and $\lim_{t\to\infty} L^{-1}[u] = 0$ (if $d = 0$, the proof stops here). Since $u, \dot{u} \in L_\infty$, it follows from Lemma 4.4 that

$$\frac{d^j}{dt^j} L^{-1}[u] \in L_\infty \text{ and is continuously differentiable} \quad \forall\, j = 1, \ldots, d$$

Since $\lim_{t\to\infty} L^{-1}[u] = 0$, $\forall \delta > 0$, $\exists T > 0$, such that $\forall t > T$, $l^{-1}[u_i] < \delta$, $\forall\, i = 1, \ldots, p$, where p is the dimension of H. Furthermore, by the mean value theorem

$$\int_t^{t+\Delta t} \frac{d}{dt} l^{-1}[u_i] = l^{-1}[u_i]|_{t+\Delta t} - l^{-1}[u_i]|_t = \Delta t\, \frac{d}{dt} l^{-1}[u_i]|_{t_i^*} \quad t \leq t_i^* \leq t + \Delta t \quad \forall\, i = 1, \ldots, p$$

so that $\forall t > T$

$$\left|\frac{d}{dt} l^{-1}[u_i]\right| \leq \frac{|l^{-1}[u_i]|_{t+\Delta t} - l^{-1}[u_i]|_t|}{\Delta t} + \Delta t \sup_{[t,t+\Delta t]} \left(\left|\frac{d^2}{dt^2} l^{-1}[u_i]\right|\right) \leq \frac{2\delta}{\Delta t} + k\, \Delta t \quad \forall\, i$$

for some constant $k > 0$, since $\frac{d^2}{dt^2} L^{-1}[u] \in L_\infty$. To prevent proliferation of constants, we will hereafter use the single symbol k to designate an arbitrarily large positive constant. If Δt is chosen $= \delta^{1/2}$, then $\forall t > T$

$$\left|\frac{d}{dt} l^{-1}[u_i]\right| \leq k\, \delta^{1/2} \quad \forall\, i$$

By a similar reasoning, $\forall t > T$,

$$\left|\frac{d^j}{dt^j} l^{-1}[u_i]\right| \leq \frac{2\left|\frac{d^{j-1}}{dt^{j-1}} l^{-1}[u_i]\right|}{\Delta t} + k\Delta t \quad \forall\, j = 1, \ldots, d \quad \forall\, i = 1, \ldots, p$$

for some constant $k > 0$, since $\frac{d^{j+1}}{dt^{j+1}} L^{-1}[u] \in L_\infty$, $\forall\, j = 1, \ldots, d$ ($\dot{u} \in L_\infty$). If Δt is chosen $= \delta^{(1/2)^j}$, then $\forall\, t > T$

$$\left|\frac{d^j}{dt^j} l^{-1}[u_i]\right| \leq k\delta^{(1/2)^j} \quad \forall\, j = 1, \ldots, d \quad \forall\, i = 1, \ldots, p$$

so that, $\forall t > T$

$$|u_i(t)| < k\delta^{(1/2)^d} + \epsilon(t) \qquad i = 1, \ldots, p$$

where $\epsilon(t)$ is an exponentially decaying term. Consequently, $\lim_{t\to\infty} u = 0$. $\square$

Bibliography

[1] G. C. Goodwin and K. S. Sin. *Adaptive Filtering, Prediction, and Control.* Prentice-Hall, Englewood Cliffs, NJ, 1984.

[2] K. J. Åström and B. Wittenmark. *Adaptive control.* Addison-Wesley, Reading, MA, 1988.

[3] S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence, and Robustness.* Prentice-Hall, Englewood Cliffs, NJ, 1989.

[4] K. S. Narendra and A. Annaswamy. *Stable adaptive systems.* Prentice-Hall, Englewood Cliffs, NJ, 1989.

[5] R. V. Monopoli and C. C. Hsing. Parameter adaptive control of multivariable systems. *Int. J. Control*, 22:313–327, 1975.

[6] H. Elliott and W. Wolovich. A parameter adaptive control structure for linear multivariable systems. *IEEE Trans. Automat. Contr.*, AC-27:340–352, 1982.

[7] M. Das. Multivariable adaptive model matching using less a priori information. *J. Dynam. Syst. Measurement Contr.*, 108:151–153, 1986.

[8] G. Tao and P. Ioannou. Robust model reference adaptive control for multivariable plants. *Int. J. Adaptive Control and Signal Processing*, 2:217–248, 1988.

[9] R. Singh and K. Narendra. A globally stable adaptive controller for multivariable systems. Technical Report 8111, Yale University, 1982.

[10] R. P. Singh and K. S. Narendra. Prior information in the design of multivariable adaptive controllers. *IEEE Trans. Automat. Contr.*, AC-29:1108–1111, 1984.

[11] H. Elliott and W. Wolovich. Parameterization issues in multivariable adaptive control. *Automatica*, 20:533–545, 1984.

[12] L. Dugard, G. C. Goodwin, and C. E. De Souza. Prior knowledge in model reference adaptive control of multinput multioutput systems. *IEEE Trans. Automat. Contr.*, AC-29:761–764, 1984.

[13] R. Johansson. Parametric models of linear multivariable systems for adaptive control. *IEEE Trans. Automat. Contr.*, AC-32:303–313, 1987.

[14] R. Ortega, J. M. Dion, F. J. Carrillo, and L. Dugard. A globally stable multivariable adaptive controller with reduced prior knowledge. In *Proc. 4th American Control Conf.*, pages 418–422, 1985.

[15] Y. S. Kim and Z. Bien. Two stable MRAC algorithms for linear discrete-time multivariable plants. *Int. J. Control*, 48:1043–1056, 1988.

[16] S. Bittanti and R. Scattolini. Multivariable self-tuning control: a model-following approach. *Int. J. Control*, 42:1035–1047, 1985.

[17] R. Scattolini and D. W. Clarke. Multivariable model-following self-tuning control with offset rejection. *Int. J. Control*, 42:1309–1322, 1985.

[18] R. Scattolini. A multivariable self-tuning controller with integral action. *Automatica*, 22:619–627, 1986.

[19] G. C. Goodwin, P. J. Ramadge, and P. E. Caines. Discrete-time multivariable adaptive control. *IEEE Trans. Automat. Contr.*, AC-25:449–456, 1980.

[20] G. C. Goodwin and R. S. Long. Generalization of results on multivariable adaptive control. *IEEE Trans. Automat. Contr.*, AC-25:1241–1245, 1980.

[21] C. A. Tsiligiannis and S. A. Svoronos. Multivariable self-tuning control via the right interactor matrix. *IEEE Trans. Automat. Contr.*, AC-31:987–989, 1986.

[22] L. Keviczky and J. Hetthessy. Self-tuning minimum variance control of MIMO discrete time systems. *Automatic Control Theory and Applications*, 5:11–17, 1977.

[23] U. Borison. Self-tuning regulators for a class of multivariable systems. *Automatica*, 15:209–215, 1979.

[24] H. N. Koivo. A multivariable self-tuning controller. *Automatica*, 16:351–366, 1980.

[25] G. C. Goodwin, P. J. Ramadge, and P. E. Caines. Discrete time stochastic adaptive control. *SIAM J. Control Optim.*, 19:829–853, 1981.

[26] L. Keviczky and K. S. P. Kumar. Multivariable self-tuning regulator with generalized cost-function. *Int. J. Control*, 33:913–921, 1981.

[27] M. M. Bayoumi, K. Y. Wong, and M. A. El-Bagoury. A self-tuning regulator for multivariable systems. *Automatica*, 17:575–592, 1981.

[28] L. Dugard, G. C. Goodwin, and X. Xianya. The role of the interactor matrix in multivariable stochastic adaptive control. *Automatica*, 20:533–545, 1984.

[29] C. Han, S. Chen, and P. D. Roberts. Modified algorithm for MIMO self-tuning control. *Int. J. Control*, 45:1515–1526, 1987.

[30] D. Praeger and P. Wellstead. Multivariable pole-assignment self-tuning regulators. *Proc. IEE*, Part D, 128:9–18, 1981.

[31] J. Mikleš. A multivariable self-tuning controller based on pole-placement design. *Automatica*, 26:293–302, 1990.

[32] H. Elliott, W. Wolovich, and M. Das. Arbitrary pole placement for linear multivariable systems. *IEEE Trans. Automat. Contr.*, AC-29:221–228, 1984.

[33] T. Djaferis, M. Das, and H. Elliott. Reduced order adaptive pole placement for multivariable systems. *IEEE Trans. Automat. Contr.*, AC-29:637–638, 1984.

[34] Y. Willner, M. Heymann, and M. Bodson. Parameter convergence and uniqueness in nonminimal parameterizations for multivariable adaptive control. *Siam J. Control Optim.*, 30:800–815, 1992.

[35] E. J. Davison. Multivariable tuning regulators: the feedforward and robust control of a general servomechanism problem. *IEEE Trans. Automat. Contr.*, AC-21:35–47, 1976.

[36] J. Penttinen and H. N. Koivo. Multivariable tuning regulators for unknown systems. *Automatica*, 16:393–398, 1980.

[37] A. H. Jones and B. Porter. Design of adaptive digital set-point tracking PID controllers incorporating recursive step-response matrix identifiers for multivariable plants. *IEEE Trans. Automat. Contr.*, AC-32:459–462, 1987.

[38] D. E. Miller and E. J. Davison. The self-tuning robust servomechanism problem. *IEEE Trans. Automat. Contr.*, AC-34:511–523, 1989.

[39] L. Dugard and J. Dion. Direct adaptive control for linear multivariable systems. *Int. J. Control*, 42:1251–1281, 1985.

[40] L. Keviczky, J. Hetthessy, M. Hilger, and J. Kolostori. Self-tuning adaptive control of cement raw material blending. *Automatica*, 14:525–532, 1978.

[41] A. J. Morris, Y. Nazer, and R. K. Wood. Multivariate self-tuning process control. *Optimal Control Applications & Methods*, 3:363–387, 1982.

[42] J. Fessl. An application of multivariable self-tuning regulators to drum boiler control. *Automatica*, 22:581–585, 1986.

[43] J. M. Martin-Sanchez and S. L. Shah. Multivariable adaptive predictive control of a binary distillation column. *Automatica*, 20:607–620, 1984.

[44] C. J. Dittmar. A hyperstable model-following flight control system used for reconfiguration following aircraft impairment. In *Proc. 7th American Control Conf.*, pages 2219–2224, 1988.

[45] W. D. Morse and K. A. Ossman. Model following reconfigurable flight control system for the AFTI/F-16' *Journal of Guidance, Control, and Dynamics*, 13:969–976, 1990.

[46] H. N. Gross and B. S. Migyanko. Application of supercontroller to fighter aircraft reconfiguration. In *Proc. 7th American Control Conf.*, pages 2232–2237, 1988.

[47] P. Chandler, M. Pachter, M. Mears, and S. Sheldon. Application of system identification techniques to flight control. In *Proc. 1st IEEE Conf. on Control Applications*, pages 611–612, 1992.

[48] D. S. Bayard, F. Y. Hadaegh, Y. Yam, R. E. Scheid, E. Mettler, and M. H. Milman. Automated on-orbit frequency domain identification for large space structures. *Automatica*, 27:931–946, 1991.

[49] D. S. Bayard. Multivariable frequency domain identification via 2-norm minimization. In *Proc. 11th American Control Conf.*, pages 1253–1257, 1992.

[50] J. M. Dion, L. Dugard, and J. Carrillo. Interactor and multivariable adaptive model matching. *IEEE Trans. Automat. Contr.*, AC-33:399–401, 1988.

[51] T. Kailath. *Linear Systems.* Prentice-Hall, Englewood Cliffs, NJ, 1980.

[52] G. O. Correa and K. Glover. Pseudo-canonical forms, identifiable parameterizations and simple parameter estimation for linear multivariable systems: input-output models. *Automatica*, 20:429–442, 1984.

[53] G. O. Correa and K. Glover. Pseudo-canonical forms, identifiable parameterizations and simple parameter estimation for linear multivariable systems: parameter estimation. *Automatica*, 20:443–452, 1984.

[54] M. Gevers and V. Wertz. Selection of identifiable parameterizations. *Control and Dynamic Systems*, 26:35–86, 1987.

[55] M. de Mathelin. Invariants and canonical forms of multivariable systems. Technical Report LASIP-90-10, Electrical & Computer Eng. Dept, Carnegie Mellon University, 1990.

[56] A. S. Morse. System invariants under feedback and cascade control. In *Lecture Notes* ; *Economics and Mathematical Systems, Vol. 131*, pages 61–74. Springer Verlag, 1976.

[57] W. A. Wolovich and P. L. Falb. Invariants and canonical forms under dynamic compensa tion. *SIAM J. Control Optim.*, 14:996–1008, 1976.

[58] V. Popov. Invariant description of linear, time-invariant, controllable systems. *SIAM J Control Optim.*, 10:252–264, 1972.

[59] G. Forney. Minimal bases of rational vector spaces with applications to multivariable linear systems. *SIAM J. Control Optim.*, 13:493–520, 1975.

[60] R. Guidorzi. Canonical structures in the identification of multivariable systems. *Automatica*, 11:361–374, 1975.

[61] S. Beghelli and R. Guidorzi. A new input-output canonical form for multivariable system. *IEEE Trans. Automat. Contr.*, AC-21:692–696, 1976.

[62] H. Elliott, R. Cristi, and M. Das. Global stability of adaptive pole placement algorithms. *IEEE Trans. Automat. Contr.*, AC-30:348–356, 1985.

[63] E. Kamen. Direct adaptive control in a state-space setting. In *Proc. 9th Int. Conf. Analysis and Optimization of Systems*, Antibes, France, 1990.

[64] D. Janecki. Stability analysis of Elliott's direct adaptive pole placement. *Systems & Control Letters*, 11:19–26, 1988.

[65] M. de Mathelin and M. Bodson. Frequency domain conditions for parameter convergence in multivariable recursive identification. *Automatica*, 26:757–767, 1990.

[66] S. Boyd and S. Sastry. Necessary and sufficient conditions for parameter convergence in adaptive control. *Automatica*, 22:629–639, 1986.

[67] M. de Mathelin and M. Bodson. Multivariable adaptive control: Prior information, parameterizations, and parameter convergence. In *Proc. 10th American Control Conf.*, pages 677–678, 1991.

[68] B.D.O. Anderson. Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. Automat. Contr.*, AC-22:83–88, 1977.

[69] G. Kreisselmeier. Stabilized least-squares type adaptive identifiers. *IEEE Trans. Automat. Contr.*, AC-33:306–310, 1990.

[70] D. R. Mudgett and A. S. Morse. Adaptive stabilization of linear systems with unknown high-frequency gains. *IEEE Trans. Automat. Contr.*, AC-30:549–554, 1985.

[71] D. R. Mudgett and A. S. Morse. Adaptive stabilization of a discrete linear system with an unknown high-frequency gain. *IEEE Trans. Automat. Contr.*, AC-30:798–799, 1985.

[72] A. S. Morse. A three-dimensional universal controller for the adaptive stabilization of any strictly proper minimum-phase system with relative degree not exceeding two. *IEEE Trans. Automat. Contr.*, AC-30:1188–1191, 1985.

[73] T. Lee and K. S. Narendra. Removing the high-frequency gain assumption in discrete adaptive control. In *Proc. 24th Conf. on Decision and Control*, pages 1198–1202, 1985.

[74] R. Lozano-Leal, J. Collado, and S. Mondié. Model reference robust adaptive control without *a priori* knowledge of the high-frequency gain. *IEEE Trans. Automat. Contr.*, AC-35:71–78, 1990.

[75] A. S. Morse, D. Q. Mayne, and G. C. Goodwin. Applications of hysteresis switching in parameter adaptive control. *IEEE Trans. Automat. Contr.*, AC-37:1343–1354, 1992.

[76] R. Lozano. Singularity-free adaptive pole-placement without resorting to persistency of excitation: detailed analysis for first order systems. *Automatica*, 28:27–33, 1992.

[77] M. de Mathelin and M. Bodson. Multivariable model reference adaptive control without constraints on the high-frequency gain matrix. In *Proc. 30th Conf. on Decision and Control*, pages 2842–2847, 1991.

[78] G. W. Stewart. *Introduction to matrix computations.* Academic Press, New-York, NY, 1973.

[79] G. W. Stewart and J.-G. Sun. *Matrix perturbation theory.* Academic Press, San Diego, CA, 1990.

[80] G. H. Golub and C. F. Van Loan. *Matrix computations.* The Johns Hopkins University Press, Baltimore, MD, 1983.

[81] C. A. Desoer and M. Vidyasagar. *Feedback systems: input-output properties.* Academic Press, New-York, NY, 1975.

[82] M. Vidyasagar. *Nonlinear Systems Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1978.

[83] V. Wertz, M. Gevers, and E. J. Hannan. The determination of optimum structures for the state space representation of multivariate stochastic processes. *IEEE Trans. Automat Contr.*, AC-27:1200–1211, 1982.

[84] A. J. M. Van Overbeek and L. Ljung. On-line structure selection for multivariable state space models. *Automatica*, 18:529–543, 1982.

[85] E. J. Hannan and L. Kavalieris. Multivariable linear time serie models. *Adv. Appl. Probab.*, 16:492–561, 1984.

[86] V. Eldem and N. Yildizbayrak. Parameter and structure identification of linear multivariable systems. *Automatica*, 24:365–373, 1988.

[87] P. Janssen, P. Stoica, T. Söderström, and P. Eykhoff. Model structure selection for multivariable systems by cross-validation methods. *Int. J. Control*, 47:1737–1758, 1988.

[88] J. J. Fuchs. Structure and order estimation of multivariable stochastic processes. *IEEE Trans. Automat. Contr.*, AC-35:1338–1341, 1990.

[89] M. de Mathelin and M. Bodson. Canonical vs pseudo-canonical forms for the structural & parametric identification of multivariable systems. In *Proc. 1st European Control Conf.*, 1991.

[90] P. Ioannou and K. Tsakalis. A robust direct adaptive controller. *IEEE Trans. Automat. Contr.*, AC-31:1033–1043, 1986.